Emergence of activation or repression in transcriptional control under a fixed molecular context

Rosa Martinez-Corral[a,b,1], Dhana Friedrich[a,f], Robert Frömel[c,d], Lars Velten[c,d], Jeremy Gunawardena[a], and Angela H. DePace[a,e,1]

[a]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.
[b]Centre for Genomic Regulation, Barcelona Collaboratorium for Modelling and Predictive Biology, 08003, Barcelona, SPAIN.
[c]Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Dr. Aiguader 88, Barcelona 08003, SPAIN.
[d]Universitat Pompeu Fabra (UPF), Barcelona, SPAIN.
[e]Howard Hughes Medical Institute
[f]Present address: Bayer AG Pharmaceuticals, 51368 Leverkusen, Germany.
[1]Correspondance: rosa.martinez@crg.eu; angela_depace@hms.harvard.edu

Abstract

For decades, studies have noted that transcription factors (TFs) can behave as either activators or repressors of different target genes. More recently, evidence suggests TFs can act on transcription simultaneously in positive and negative ways. Here we use biophysical models of gene regulation to define, conceptualize and explore these two aspects of TF action: "duality", where TFs can be overall both activators and repressors at the level of the transcriptional response, and "coherent and incoherent" modes of regulation, where TFs act mechanistically on a given target gene either as an activator or a repressor (coherent) or as both (incoherent). For incoherent TFs, the overall response depends on three kinds of features: the TF's mechanistic effects, the dynamics and effects of additional regulatory molecules or the transcriptional machinery, and the occupancy of the TF on DNA. Therefore, activation or repression can be tuned by just the TF-DNA binding affinity, or the number of TF binding sites, given an otherwise fixed molecular context. Moreover, incoherent TFs can cause non-monotonic transcriptional responses, increasing over a certain concentration range and decreasing outside the range, and we clarify the relationship between non-monotonicity and common assumptions of gene regulation models. Using the mammalian SP1 as a case study and well controlled, synthetically designed target sequences, we find experimental evidence for incoherent action and activation, repression or non-monotonicity tuned by affinity. Our work highlights the importance of moving from a TF-centric view to a systems view when reasoning about transcriptional control.

Keywords: transcription factor | duality | non-monotonicity | gene regulation | non-equilibrium regulation

1 Introduction

Transcription factors (TFs) are proteins that bind to DNA sequences to modulate transcription of target genes. It is common to consider TFs as either activators or repressors based on the effect that increasing the concentration of a TF has on target gene expression: increasing the concentration of an activator leads to a rise in transcription, whereas increasing the concentration of a repressor reduces it.

Despite this binary classification, it is well known that some TFs are "dual" or "bifunctional" and behave as either activators or repressors [1, 2, 3, 4]. For example, Dorsal, a Drosophila TF of the NF-kB family, can activate target genes on its own and can repress when it co-binds with other regulators with which it collaborates to recruit co-repressors [5, 6, 7, 8]. Similar explanations for duality based on a change in the molecular partners of a TF, conformation, post-translational state or other features relevant to its mechanistic functioning have been given for the Drosophila TF Krüppel [9, 10, 11, 12], and many mammalian TFs, including the glucocorticoid receptor [13], FOXO3 [14], Sp3 [15], Myc [16], Yin-Yang 1 [17] and NF-κB [18, 19].

Experimental data also suggests that TFs can simultaneously activate and repress a given target gene. For example, Bicoid, a classical *Drosophila* activator, was found to both promote and prevent progression of a gene regulatory system into transcriptionally active states [20]. In line with these observations, Bicoid had been suggested to interact with both histone acetyltransferases (commonly activating) and histone deacetylases (commonly repressive) [21], and it was found to exhibit increased activity upon deletion of a fragment of the protein, which was considered to have a repressive effect [22, 23]. Similarly, mammalian TFs with both activating and repressive domains can be found, as described for RFX1 [24], Oct3/4 [25] and some proteins of the KRAB family [26]. Recent experiments have also found domains with dual activities [27, 28, 29]. Even in bacteria, the prokaryotic TF CpxR was found to have simultaneously positive and negative effects on transcription [30].

Here we introduce the term "incoherent" to describe this latter form of gene regulation. Our goal is to distinguish modes of action for a dual TF. For example, a dual TF could act coherently (e.g. as an activator at some targets and as a repressor at others). Or a dual TF could act incoherently (e.g. as both an activator and a repressor at a given target gene). We use mathematical models of gene regulation to conceptualise both kinds of modes of action for dual TFs (coherent and incoherent), clarify how each mode relates to the TF's overall effect (activation or repression), and determine what can tune the overall effect between activation and repression.

We distinguish three categories of features, which correspond to different sets of parameters in our models. First, we consider parameters related to the effects of the TF on the proteins involved in the transcription process. We will refer to this type of effect as "the mechanistic effect" of the TF, which could be given by the TFs' intrinsic physicochemical properties, or its direct or indirect interactions with coregulators and the transcriptional machinery. We remain agnostic to the exact molecular implementation, and focus just on the implications of the coherent or incoherent modes of action, and different regulatory strengths within the incoherent mode. Commonly, switching between the two types of coherent modes (activating and repressing) has been invoked to reason about duality. More recently, Guharajan *et al.* showed that the overall response to a TF could depend on the location of its binding site, and these data were explained by a model where the binding location determined the mode (coherent or incoherent) and strength of the TFs' effects in each mode [30]. Second, we consider parameters that are not directly related to the mechanistic effect of the TF itself, but that modulate background regulatory processes, such as basal chromatin dynamics and promoter strength. Ali *et al.* showed theoretically that the promoter strength can tune the overall effect between activation and repression for an incoherent TF [31]. Finally, we consider the effect of parameters that correspond to the TF-DNA binding kinetics or affinity. We show that these can also tune the overall effect between activation and repression, which to our knowledge has not been explicitly described before.

In addition to monotonic activation or repression, responses can be non-monotonic, where the response increases with TF concentration over a concentration range and decreases outside that range. Exper-

imentally, non-monotonicity in the effects of a TF has been explained in terms of squelching [32, 33, 34]: at high concentration, the TF sequesters the machinery necessary for transcriptional activation away from the target genes, which causes the observed decrease not only at the gene of interest being assessed but also at other genes. Other explanations invoke stress-related effects (E.g. [29]). However, non-monotonicity can also reflect the regulatory effects of a TF acting in incoherent mode. Gedeon *et al.* investigated the emergence of non-monotonicity using thermodynamic models of gene regulation, where TFs are assumed to recruit polymerase under conditions of thermodynamic equilibrium [35]. In this setting, for a TF binding to a single site, the responses were monotonic, whereas the responses could become non-monotonic for multiple sites. More recently, non–monotonic responses have been found to arise in single-site models when polymerase recruitment is considered away from thermodynamic equilibrium [36, 31]. These results suggest that non-monotonicity could be a signature of non-equilibrium regulated recruitment, a topic of active research [37, 38, 39, 40].

Here we investigate the relationship between TF regulatory mode, whether or not regulation and polymerase recruitment occur at equilibrium, and the emergence of non-monotonic responses. We model gene regulation using the linear framework [41, 37], which allows us to study both equilibrium and non-equilibrium steady-state systems with the same mathematical formalism and perform certain analytical calculations more easily than with other methods. We interrogate two models of gene regulation: a model of regulated recruitment of RNA Polymerase, and a model of the regulation of the RNA Polymerase transcriptional cycle. This latter model accounts for the intrinsically dissipative nature of the transcription process, while allowing for the binding of TFs to take place at thermodynamic equilibrium [40]. In this case, we find that non-monotonic responses can also arise even for a TF binding to a single site, which is not possible for single site models of equilibrium regulated polymerase recruitment. Therefore, we clarify that in the absence of squelching or stress, non-monotonic transcriptional responses reflect incoherent regulation, but not necessarily non-equilibrium regulated polymerase recruitment.

In order to illustrate the experimental significance of these findings, we investigate the effect of the mammalian TF SP1 on synthetic regulatory sequences. We find evidence for non-monotonicity and an affinity-dependent change in the overall effect of the response. We interpret these findings as an indication of incoherent duality that arises in the absence of changes to co-regulatory proteins or to the TFs interactions with the transcriptional machinery.

2 Results

2.1 A model of regulated recruitment

We begin by considering the well-established recruitment view of transcriptional control [42], where the TF is assumed to modulate the recruitment of the transcriptional machinery to the gene promoter. We consider a TF that binds to only one regulatory site, and that modulates transcription through two molecular processes (Fig. 1A). First, we assume that the transcription start site can exist in two conformations, for which polymerase has different affinity, and the TF can modulate the probability of each conformation. The two conformations are interpreted here as the region being either occupied by a nucleosome or free [43, 44], and we will refer to them as closed or open, respectively, although other interpretations are possible (see Discussion). Second, the TF can directly modulate polymerase binding through binding cooperativity at a given conformation.

Formally, we model this system using the linear framework [41, 45, 46], a graph-based approach to analyse Markov Processes (Methods). We consider that a gene regulatory region can be in a series of states. In the current model, the states are determined by whether TF and polymerase are bound or not, and which of the two conformations the system is in. We assume that transitions between states are constant over time and follow Markovian dynamics. Such system can be represented by a graph (linear framework graph, Fig. 1E, see also Fig. S1), with the vertices corresponding to the states of the system and edge labels the transition rates, in dimensions of inverse time.
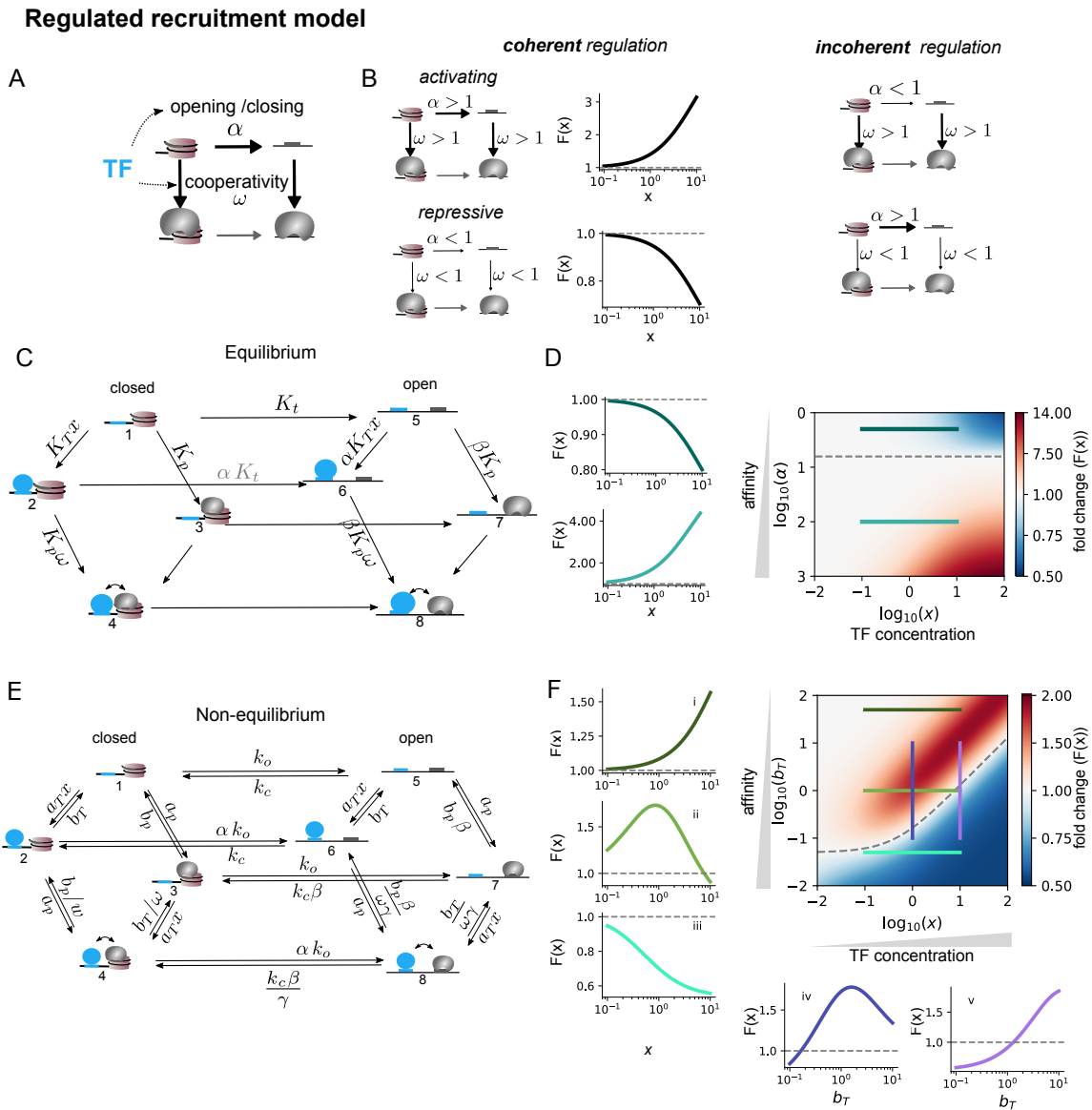
Figure 1: Duality in a model of regulated recruitment. A) Schema of the model. See text for details. Gray shape: Polymerase. Cylinder: nucleosome. B) Parametric regimes corresponding to coherent and incoherent regulation. The lineplots correspond to the model in panel C, for the same parameters in D except that $\alpha = 10, \omega = 2$ for activation, and $\alpha = 0.01, \omega = 0.5$ for repression. C)Equilibrium graph of the system. Blue disc: TF. The edges denote reversible transitions, with the labels given by the ratios between forward and backward transition rates. K_T is the TF binding affinity to the closed conformation. x is TF concentration. K_p is the product of Polymerase concentration and its binding affinity for the closed conformation. In the absence of TF and polymerase, the ratio of the transition rate from the closed to the open conformation over its reverse is given by K_t. We consider $K_t < 1$, so that the closed conformation is favored for the unbound state. Polymerase binding constant to the open conformation is increased by a factor $\beta > 1$. ω denotes the binding cooperativity at each conformation. D) Monotonic activation or repression for the model in panel C in the incoherent regime ($\alpha > 1, \omega = 0.5$) as a function of α. B,D (a.u.): $K_T = 0.1, K_p = 0.01, K_t = 0.005, \beta = 50, q_3 = q_4 = q_7 = q_8 = 1$. E) Non-equilibrium counterpart of the graph in C. See text and SI for details. F) Example behaviour of the model in E for the TF in incoherent mode. Panels i-iii show the response to TF concentration x at three unbinding rates (horizontal lines on the colormap, color-coded). Panels iv-v correspond to two fixed concentrations and show the response to varying unbinding rate b_T (vertical lines on the colormap, color-coded). Parameter values (a.u.): $a_x = 1, b_x = 1, a_p = 0.1, b_p = 100, k_o = 0.01, k_c = 0.5, \beta = 0.02$.

4

We are interested in the steady-state transcription rate $r^*(x)$, which is assumed to be a linear combination of the steady-state probability distribution of the n system states:

$$r^*(x) = \sum_{i=1}^{n} q_i P_i^*(x) \tag{1}$$

where $P_i^*(x)$ is the steady-state probability of vertex i at TF concentration x, and q_i is the transcription rate from that state. By assuming linear mRNA decay at rate δ, at steady state the mRNA concentration is given by

$$m^*(x) = \frac{r^*(x)}{\delta} \tag{2}$$

To characterise the effect of the TF, we consider the fold change in mRNA levels $F(x)$, which is equivalent to the ratio of the steady-state transcription rate under a given TF concentration x, to the basal transcription rate (at $x = 0$):

$$F(x) = \frac{r^*(x)/\delta}{r^*(0)/\delta} = \frac{r^*(x)}{r^*(0)} \tag{3}$$

We are interested in understanding how does the TF's mode of action relate to the curve $r^*(x)$, and equivalently, $F(x)$. In particular, under what parametric conditions is $dF(x)/dx > 0$ (activation) or $dF(x)/dx < 0$ (repression) and what can cause a switch between the two overall effects?

2.1.1 Monotonic duality at thermodynamic equilibrium

Following the classical thermodynamic models of gene regulation [47, 48], we first consider the model under assumptions of thermodynamic equilibrium (Fig. 1C), where the steady state probability distribution of the system is determined by the free energies of the various states, in the absence of external sources of energy. Under equilibrium conditions, for the purpose of analysing the steady state, we can replace linear framework graphs, which have forward and backward transitions among states (Fig. 1E), by equilibrium graphs (Fig. 1C), where we only specify transitions in one direction (despite they are reversible) with the edge labels being the ratios between the forward and the backward transitions. Accordingly, in the graph of Fig. 1C we only show edges in one direction, with the corresponding parameters being rate ratios. Each polymerase-bound state is assumed to have a non-zero transcription rate q_i, assumed all equal for now (see SI Appendix for a discussion of different q_i), so that the overall steady-state transcription rate is given by:

$$r^*(x) = q(P_3^*(x) + P_4^*(x) + P_7^*(x) + P_8^*(x)) \tag{4}$$

We assume that polymerase has a higher affinity for the open conformation ($\beta > 1$). We use α to parameterise the opening/closing effect of the TF, with $\alpha > 1$ corresponding to opening, and $\alpha < 1$ to closing. So, the value of 1 represents a transition point between two different mechanistic regimes: a regime where the TF promotes binding ($\alpha > 1$) and one where it disfavors it ($\alpha < 1$). The second effect of the TF is assumed to correspond to binding cooperativity at a given conformation (ω), such that when the TF is bound, polymerase binding is favored or disfavored. Again, the value of 1 is the limit between two regimes of molecular mechanism, with $\omega > 1$ corresponding to the TF enhancing binding, and $\omega < 1$ to reduced binding. Note that due to the equilibrium assumption, this cooperativity is reciprocal, so that polymerase also favors/disfavors TF binding. Interactions with different cofactors, or different binding site locations, may determine one or another regime, with $\omega < 1$ for example corresponding to a bind-

ing site occluding the transcription start site. Regardless of the exact molecular implementation, in this model, coherent regulation will correspond to both α and ω greater or smaller than 1, and incoherent regulation to having one value above one and the other below one (Fig. 1B). (More complicated scenarios are considered in the SI Appendix).

Coherent regulation: If the TF promotes opening ($\alpha > 1$) and polymerase binding through positive binding cooperativity ($\omega > 1$), then both α and ω coherently enhance polymerase binding (since we assume its binding is favored in the open conformation). As expected, the TF causes activation, so that a higher TF concentration leads to a higher fold change in the response (Fig. 1B-top). In contrast, if the TF disfavors the open state ($\alpha < 1$) and exhibits negative binding cooperativity with polymerase ($\omega < 1$), then it causes repression (Fig. 1B-bottom). This is analytically proven in the SI Appendix, where the case where the transcription rates are different for the different states is also considered. For the response to switch between activation and repression, there must be a change in the TF mechanism of action, switching from enhancing transcription-promoting processes to repressive processes, in line with the common understanding of duality.

Incoherent regulation: We now consider the "incoherent" mode of TF regulation, with $\alpha > 1$ and $\omega < 1$ (or vice-versa). A potential molecular implementation of this regime would be the bound TF competing with the nucleosome ($\alpha > 1$) but interfering with the assembly of the polymerase complex ($\omega < 1$). In this case, the overall effect of the TF depends on the balance between the different transitions of the system, as encoded by the parameter values. As expected, the overall effect may be tuned by the strength of the positive and negative effects of the TF and the balance between the two. We do not focus on this, since this corresponds to the common view of duality as being dependent on changes to how the TF acts mechanistically. Rather, we are interested in how the overall effect can change between activation and repression even if the TF mechanistic effects remain constant.

On the one hand, activation or repression may depend on the values of the parameters that account for the dynamics and concentrations of molecules other than the TF. For example, K_t determines the distribution of the two conformations in the absence of TF and may be interpreted to depend on the concentration of some chromatin regulator. We show in Fig. S2A that just varying K_t can cause the "incoherent" TF to change from behaving as an activator to behaving as a repressor. Similarly, Ali et al. [31] showed that the promoter strength (which would correspond to βK_p in our model) could be another feature capable of tuning the response of an incoherently acting TF. Therefore, a small change in the dynamics of the system may be enough to tilt the balance between activation and repression, even if the same molecules are involved in the TFs' response and its direct effects on the transcription process remain unchanged.

On the other hand, we may now focus on the interaction between the TF and the DNA. Due to the equilibrium constraints, α, the parameter that regulates the effect of the TF on the distribution of conformations, must equal the factor change in affinity of the TF for the open conformation relative to that for the closed conformation. For $\alpha > 1$, α can be considered as being related to the affinity of the DNA binding site, with higher affinity sites corresponding to higher α values and more effective competition between the TF and the nucleosome. As shown in the colormap in Fig. 1D, for some values of $\omega < 1$, changing the value of α despite maintaining it above 1 (the TF always promotes the open conformation) can cause the TF to switch from behaving as a repressor at low values of α to behaving as an activator at high values of α. Therefore, this interpretation of α as being related to the TF-DNA affinity suggests that changes in the DNA binding affinity of the TF can cause the response to switch between repression and activation, with no alterations to the mechanistic functioning of the protein or molecular partners. A clearer effect of the TF-DNA binding affinity is discussed in the next subsection, where we consider the model away from equilibrium, so that the opening effects can be uncoupled from the affinity.

The response to changing TF concentration is always monotonic under this equilibrium model for a single regulatory site (SI Appendix). In fact, irrespective of the complexity of the model, any equilibrium model that assumes a TF regulates a gene by binding to a single site, with the free TF concentration not reduced by binding, and with the response a linear combination of the steady-state probabilities of the various states, will lead to a monotonic response to changing TF concentration (SI Appendix),

in line with [35]. We show in the next section that this is no longer the case when the assumption of thermodynamic equilibrium is dropped, and we show subsequently in the paper that non-monotonicity may also appear when considering the effects of a TF on the dissipative transcription cycle, even under equilibrium binding conditions.

2.1.2 Non-monotonic duality away from equilibrium

In the previous section, the equilibrium assumption would be appropriate to model a TF that displaces nucleosomes by taking advantage of nucleosome fluctuations and/or by having an intrinsic ability to bind closed chromatin and destabilize nucleosomes. This has been well studied for the so-called "pioneer factors" *in vitro* [49, 50]. However, there is increasing evidence that *in vivo* the effects of TFs on chromatin are intimately coupled to energy dissipative cofactors [**Iwafuchi-Doi2019-ik**, 51], including chromatin remodellers that hydrolyze ATP [51, 52], and/or histone modifying enzymes [53]. Recent experimental evidence suggests that these active processes may be continuously required to maintain gene expression by these factors [54, 52], suggesting it may be more appropriate to consider regulatory systems away from thermodynamic equilibrium. This situation can be accounted by the model in Fig. 1E, where certain assumptions reduce the number of free parameters in the system so that it is simpler to reason about the mechanism of action of the TF (details are given in SI Appendix, as well as a more general case. See also Fig. S4). Contrary to the equilibrium case, the TF's affinity is now parameterised as being independent on the conformation, and independent on the opening strength α. The TF is also assumed to affect polymerase recruitment through affecting its dissociation rate by a factor ω.

As in the equilibrium counterpart, monotonic activation or repression correspond, respectively, to the TF coherently enhancing polymerase binding by both mechanisms ($\alpha > 1, \omega > 1$), or reducing it ($\alpha < 1, \omega < 1$) (See SI for monotonicity proof.) In the incoherent regulatory mode, the TF enhances the open conformation ($\alpha > 1$) but reduces polymerase binding by negative cooperativity ($\omega < 1$) (or vice-versa). As in the equilibrium counterpart, increasing the TF concentration can either increase or decrease expression depending on the parameter values, and the overall effect can depend on parameters not directly related to the TF mechanistic effect (Fig. S3A). As before, activation or repression may also depend just on the TF-DNA binding affinity, in this case assumed to be determined by the unbinding rate (Fig. 1F).

Notably, in contrast to the equilibrium counterpart and the non-equilibrium coherent regulatory mode, now non-monotonic responses are also possible under some parameter combinations. Fig. 1F shows the results of an example parameter set (see also Fig. S3). The colormap shows that the overall TF effect is determined nontrivially by the interplay between the TF-DNA binding affinity (modelled as being determined by the TF unbinding rate b_T) and the TF concentration. For a range of unbinding rates, the effect of increasing TF concentration is non-monotonic, with the TF first increasing expression at relatively low concentrations, and then reducing expression at relatively high concentrations (colormap and plot ii). If we focus our attention on a fixed, intermediate concentration range, then depending on the unbinding rate the TF appears to behave as either an activator or a repressor, as shown on the example plots on the left of the colormap. (Fig. 1F,i-iii). Similarly, the TF concentration can tune the response to unbinding rate (iv,v).

We noticed that with the simplifications considered here, including considering the expression rates q_i all equal and equal cooperativities across conformations ($\gamma = 1$), the non-monotonic responses were always bell-shaped, as in Fig. 1F-ii, iv. However, when these simplifications were lifted, which corresponds to assuming that there are more parameters through which the TF can act and which contribute to defining coherent or incoherent regulation, then we observed that U-shaped responses could also arise (Fig. S3B-D).

2.2 Duality when the TF regulates the transcription cycle

In the previous sections, we have analysed models that focus on the regulation of RNA polymerase recruitment. However, it is well-known that regulation can also occur downstream, for example at the level of pause-release and elongation [55]. These processes can be conceptualised in terms of a transcriptional cycle (Fig. 2).

For simplicity, we consider a cycle with three states, as in [56] (Fig. 2A). The three states of the cycle can be interpreted to represent unbound polymerase (state 1), bound but inactive polymerase (state 2), and actively transcribing polymerase (state 3). The first transition is assumed to be reversible, whereas the other two are modelled as irreversible in agreement with the macroscopic irreversibility of these processes in physiological scenarios.

In the first model we consider (Fig. 2C), the TF is assumed to bind to and unbind from all three polymerase cycle states with the same kinetics, parameterized by a binding rate a_T and an unbinding rate b_T. Moreover, the TF is assumed to affect the transitions in the polymerase cycle only while it is bound. This is modelled by assigning a different parameter value for a transition rate when the TF is bound as compared to the basal one, with the fold change in the parameter value given by $\epsilon_i > 0$. If the TF enhances (reduces) a given rate, ϵ_i is > 1 (< 1) when the TF is bound. The transcription rate is assumed to be proportional to the transitions from state 3 to state 1 [56]:

$$r(x) = k_3 P_{3,\emptyset} + \epsilon_3 k_3 P_{3,T} \tag{5}$$

Depending on the effects of the TF on the rates, given by the ϵ_i parameters, we can distinguish three modes of TF action or parametric regimes (Fig. 2B). If the TF acts coherently enhancing the transcription cycle, by either increasing or not modifying the clockwise rates, or decreasing or not modifying the counterclockwise rate, then the TF will behave as an activator if at least it affects one rate. Conversely, if it coherently acts negatively on the cycle, it behaves as a repressor (See SI Appendix for proof).

The third regime, incoherent mode, corresponds to the TF modulating at least two steps in the cycle but in opposite directions. For example, we may assume it increases the transition from state 1 to 2 ($\epsilon_1 > 1$) and decreases that from state 2 to 3 ($\epsilon_2 < 1$). In this case, the balance between the different parameters ultimately determines the overall effect of the TF as its concentration is increased, and as before, non-monotonic responses are possible. Again, changing the TF binding affinity can modulate its overall effect: over a limited concentration range, changing the unbinding rate of the TF (b_T) can mediate the transition from one to another response type. In the example of Fig. 2C, the colormap and left plots show that the TF acts as a repressor for low unbinding rate, activator for high unbinding rate, or produces a non-monotonic output for intermediate unbinding rate values. Similarly, these different overall effects are observed for fixed TF concentrations, as the response is evaluated as a function of the unbinding rate (Fig. 2C-bottom panels). We confirmed that the non-monotonicity is not a consequence of the specific details of this model, but also holds under variations of it with the reverse order of the transitions affected positively and negatively, other points of control of the TF, reversibility patterns, and/or more states (SI Appendix, Fig. S5). For a model with 5 states, we also found U-shaped responses under some settings, which we did not find for any of the variations tested for the 3 state model.

Overall, the models considered in Figs. 1-2 suggest that irrespective of the exact molecular implementation, when a TF regulates more than one process in transcription, we can distinguish three regimes, or modes of TF molecular action. The first two regimes, coherent activation and repression, lead to monotonically increasing or decreasing responses, respectively, and a change in the TF mode of action is required to switch between the two. The third regime is the incoherent regulatory mode. In this case, the response can be tuned between activation and repression by parameters that affect the dynamics of the system, in the absence of a switch in the TF mode of action, and the response can be non-monotonic under non-equilibrium, kinetic models.
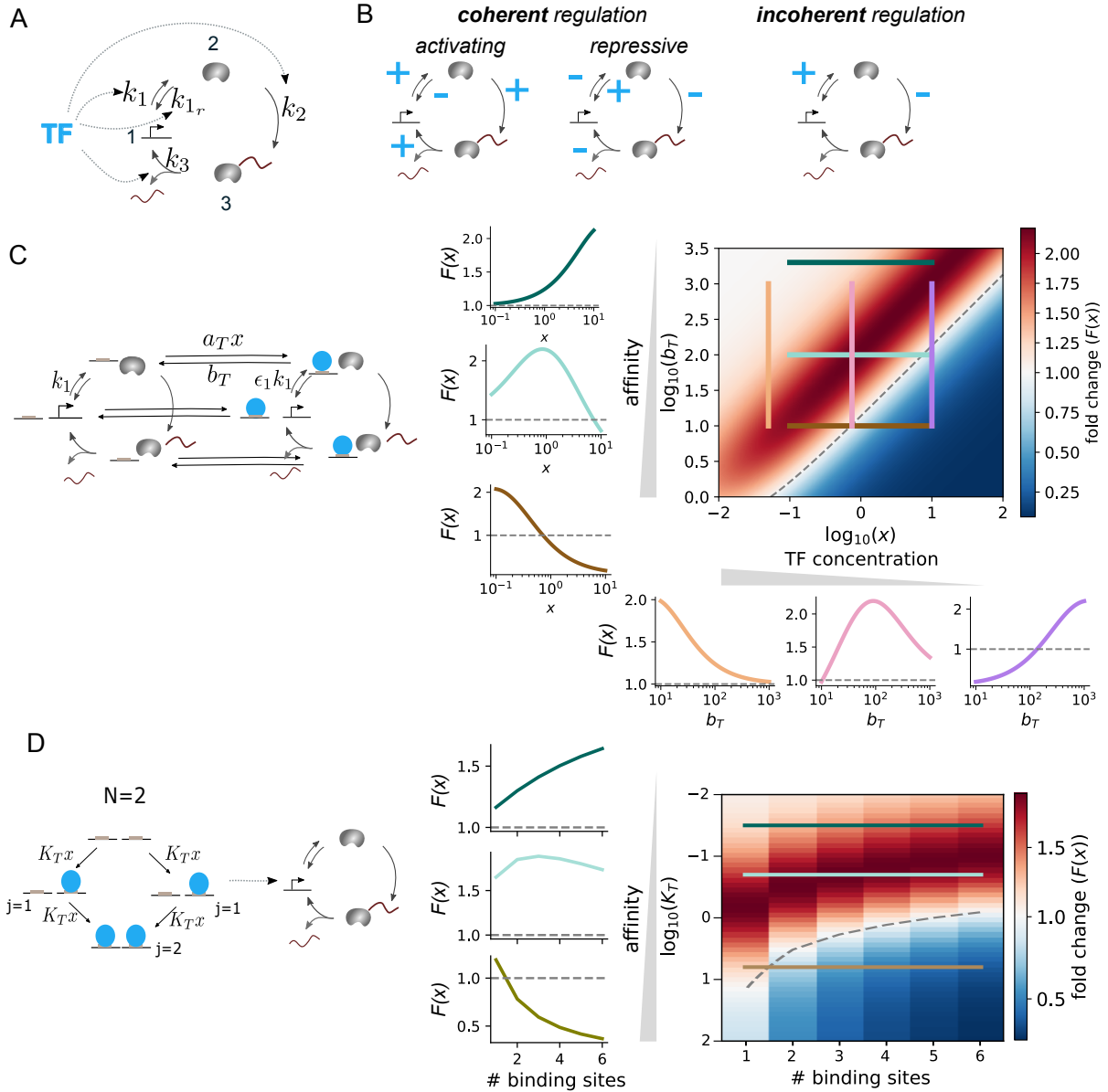
Transcription cycle model

Figure 2: Duality when a TF regulates the RNA polymerase transcriptional cycle. A) Schema of the model: a TF regulates one or multiple transitions of a three-state transcriptional cycle. B) Example of regulatory effects corresponding to coherent or incoherent regulation. A plus sign means that the TF accelerates that transition, a minus sign means it decelerates it. See text for details. C) Left: Graph of the system where TF affects a rate while bound. Only specific edge labels are shown for clarity. See text for details. D) Model where the transcription cycle rates are modulated by the equilibrium average number of bound TF molecules. The graph exemplifies the case of independent binding to N=2 sites. The colormap shows the fold change as a function of affinity and number of binding sites, with the plots on the left showing fold change as a function of number of sites for three different affinity values K_T (0.03, 0.2, 6.31). Parameter values (a.u.): $k_{1,0} = 0.5$, $k_{1,sat} = 50$, $s_1 = 0.2$, $h_1 = 1$, $k_{1r} = 50$, $k_{2,0} = 0.5$, $k_{2,sat} = 0.005$, $s_2 = 0.1$, $h_2 = 1$, $k_3 = 0.1$, $K_T = 0.174$. See text for details. See also Fig. S6.

9

2.3 Non-monotonicity with equilibrium binding

In the past section, we have assumed that the TF has an effect on the rates of the transcriptional cycle while it is bound. If we consider that the transition from state 1 and 2 represents polymerase recruitment, and $\epsilon_1 \neq 1$ or $\epsilon_{1r} \neq 1$, then the cycle that involves TF binding and polymerase recruitment does not obey detailed balance, representing non-equilibrium behaviour that involves the binding of the TF and polymerase recruitment. An alternative common in the literature is to consider a separation of timescales in the system where TF binding is assumed to be sufficiently fast that it reaches thermodynamic equilibrium, and the effects are a function of the equilibrium occupancy of the TF on the DNA. So, we consider a last model where we adapt the previous transcription cycle model to this scenario: the average number of TF molecules bound to the DNA, given by $a = \frac{K_T x}{1 + K_T x}$ (for a single site with affinity K_T), affects a given rate i from a basal value $k_{i,0}$ to a new value k_i (Fig. 2D). We assume that the effect can be non-linear and saturable at high occupancies:

$$k_i = k_{i,0} + (\epsilon_i)\left(\frac{a^h}{s_i^h + a^h}\right) \tag{6}$$

with $\epsilon_i = k_{i,sat} - k_{i,0}$ and $k_{i,sat} > 0$ a theoretical maximum or minimum rate value attainable under that TF. We find that even in this case non-monotonic responses are also possible when considering incoherent TFs (e.g $\epsilon_1 > 0$ and $\epsilon_2 < 0$ in the example of Fig. S6A), even with $h = 1$, although we tended to find more pronounced non-monotonicity over a shorter TF concentration range with higher h values (Fig. S6B).

2.4 Multiple sites

In the previous section, the effects of increasing TF concentration result from increased occupancy of the TF on the DNA (a). Thus, we tested whether the same behaviour could be obtained when increasing occupancy by keeping the TF concentration fixed and increasing the number of binding sites (assuming the TF mechanism is not dependent on binding site location). The equilibrium average number of TF molecules bound is given by $a = \sum_{j=0}^{N} (j \sum_{\#\nu=j} P_\nu^*)$, where the internal sum runs over all graph vertices with j molecules bound (Fig. 2D), and the external sum runs over the number of sites N. We assume all sites have the same affinity, and binding is independent among them (no binding cooperativity). Also in this case, depending on the affinity of the sites, increasing site number can lead to activation or repression, and non-monotonicity with respect to the number of sites (Fig. 2D-right). Similar results can be obtained for the model of Fig. 2C, when extended to have multiple TF binding sites and assuming that the effect of a TF on a rate increases linearly with the number of sites bound, irrespective of the identity of the sites (Fig. S7).

2.5 Experimental evidence for non-monotonicity and affinity-dependent activation or repression

As mentioned in the Introduction, experimentally, non-monotonicity in the effects of a TF has commonly been explained in terms of squelching [32, 33, 34], stress (E.g. [29]) or difficulties with TF overexpression. Accordingly, during preliminary work for a previous project [56], we attributed to these some non-monotonic input-output responses we observed. However, the theoretical results presented above suggest an alternative, where non-monotonicity reflects underlying TF incoherent regulation. In the light of this, we decided to reexamine the behaviour of this previous experimental system.

The system is based on a synthetic SP1 TF fusion acting on a reporter (Fig. 3A). More specifically, a domain of SP1 was fused to an artificially-designed zinc-finger DNA binding domain that binds to a (single) 20-bp artificial binding site that does not exist in the mammalian genome, upstream of a CMV promoter that regulates an eGFP reporter. The reporter construct was genomically integrated in a Hek293T cell line [56]. Input TF concentration is controlled by transfection (Fig. S9B), [56], and the transcriptional
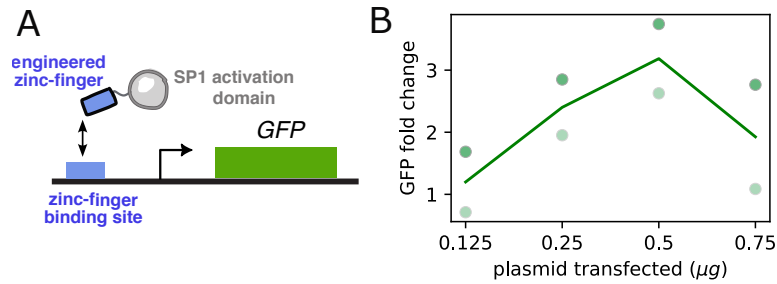
Figure 3: Non-monotonicity in the response of a reporter regulated by a synthetic ZF-SP1 TF that binds to a single regulatory site. A) Cartoon of the experimental setup. See text for details. B) GFP mRNA fold change, as measured by qPCR, as a function of amount of input TF-encoding plasmids transfected. The line represents the mean of two biological repeats, with the individual datapoints given by the dots, where each shade corresponds to an experiment.

response by monitoring reporter mRNA levels by qRT-PCR. We first selected a range of synTF plasmid concentration over which increases in input TF does not affect the expression of GAPDH nor p21 (Fig. S9A), which would suggest neither squelching (GAPDH) nor stress (p21) is relevant in this concentration range. Despite this, we could see a non-monotonic response in the reporter (Fig. 3B). Together with our modelling work, this suggests that non-monotonicity due to incoherent regulation can arise in this experimental system.

In order to further rule out the effect of squelching, and assess the significance of incoherent regulation for endogenous SP1, we turned to a massively parallel reporter assay approach. Here, we assess the effect of varying binding site number and/or affinity, parameters that our models show can also tune the overall TF effect, while TF concentration remains at its endogenous level. Thus, the effects should not be due to squelching. We performed a lentiviral massively parallel reporter assay (lentiMPRA, [57]) in K562 cells, with a library of 276 synthetic regulatory sequences combining 1-6 SP1 binding sites and 6 affinity ranges (such that all sites in a design corresponded to the same affinity range). We included different orientations and spacings of sites and used random background DNA, in order to reduce systematic biases. Binding sites were placed upstream of a minimal promoter driving a barcode and GFP expression (Fig. 4A, Methods). Activity was quantified as the logarithm of the ratio of RNA-level barcode codes over DNA-level barcode codes for each candidate regulatory element. Comparison of two technical replicates gave a high correlation (Fig. 4B), demonstrating a good quality of the data. We further normalised the activity measure by subtracting the mean activity of regulatory elements with only random sequences, so that values above or below 0 correspond to more or less expression as compared to this average activity of random DNA.

When we plotted (normalised) activity as a function of number of binding sites, we saw that the average of the data exhibited an affinity-dependent switch, with an activating trend for lower affinities, non-monotonicity for intermediate affinities and repression for higher affinities (Fig. 4D-E), in line with the modelling results in Fig. 2D. These trends were supported by the fits of the data with generalised additive models under shape constraints based on P-splines [58](Fig. S9C). Overall, we also saw a non-monotonic response (best fitted by a model with concave constraint) when we combined the number of sites and affinity range of a sequence into a single score ("sum of affinities", Methods), with the minimum score corresponding to one, low-affinity site, and the maximum corresponding to 6, high affinity sites (Fig. 4C). These data agree well with the behaviour of our models, which we interpret as an indication that the affinity-dependent switch and non-monotonicity arising from incoherent regulation exhibited by the models is experimentally significant for reasoning about the effects of endogenous TFs.
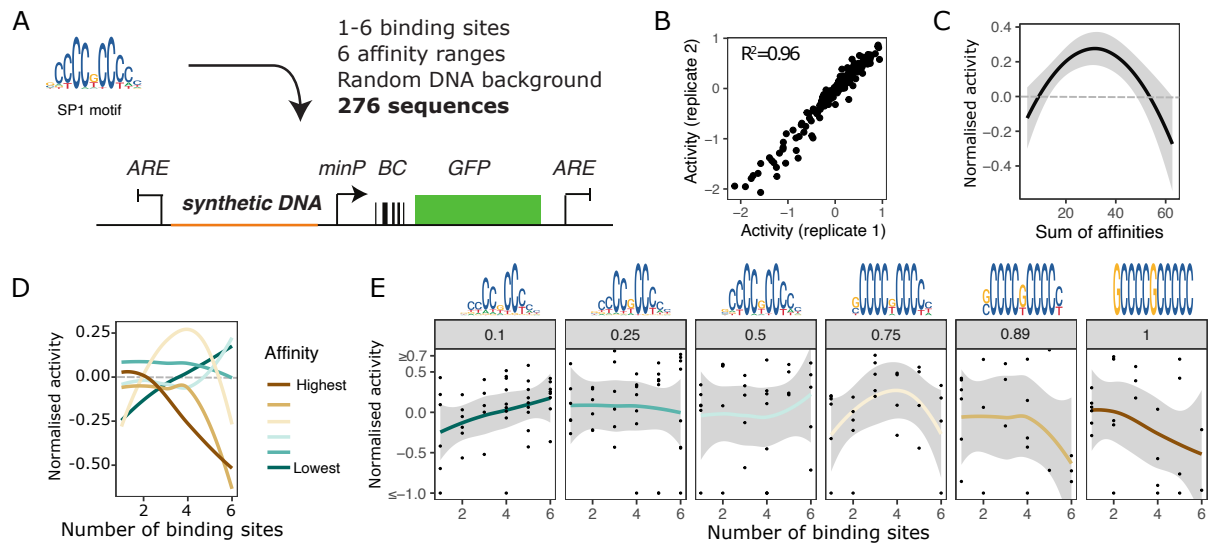
11

Figure 4: Non-monotonicity and affinity-dependent activation or repression by SP1 in K562 cells by lenti-MPRA. A) Cartoon of the experimental setup. See text and [57] for details. minP: minimal promoter. BC: barcode. ARE: anti-repressor element. B) Correlation of raw activity measures among two technical replicates. C) Mean activity as a function of the sum of affinities score (Methods). D-E) Normalised activity as a function of number of binding sites, by affinity (Methods). D) Mean activity, averaging over the data corresponding to different orientations, spacings and random DNA around the sites. E) Mean activity alongside individual datapoints (corresponding to different orientations, spacings and random DNA). In C, D, E, the average data lines were plotted with the geomsmooth function of ggplot, with the error representing the 95% confidence interval of the mean.

3 Discussion

For decades, the effects of TFs have been studied through genetic manipulation of TF levels or overexpression experiments. In these experiments, many TFs have been observed to be dual, acting as activators in some genes or cell states and repressors in others. This behaviour has commonly been interpreted under the assumption that TFs act coherently at the level of their interactions with the transcriptional machinery, such that a change in the overall effect on transcription must correspond to a change in the mechanistic effect of the regulating TF. However, increasing evidence for TFs acting incoherently requires us to revisit this assumption. In this work, we aimed to provide a conceptualisation of TF duality and to clarify how overall activation or repression at the level of the transcriptional response to a TF relate to the regulatory interactions occurring at that particular gene.

We note that the mechanistic effects of TFs can be achieved in multiple ways, and we have therefore considered mechanistic effects at a coarse level, remaining agnostic as to their exact molecular implementation. TF effects could be determined by intrinsic physicochemical properties of the TF molecule or through interactions with other molecular partners. To illustrate one possibility, in SI Appendix section *Incoherent regulation through functional interference among activators* (Fig. S8) we discuss a potential mechanism that may cause the simultaneous positive and negative effects of a TF as a result of interference with another regulatory molecule acting on the same gene. In addition, we note that our models might be interpreted slightly differently than in our explanations in the main text. For example, the two conformations of the regulated recruitment model could be related to Mediator binding [38].

Regardless of the exact molecular implementation, we have conceptualised two modes of TF mechanism: coherent and incoherent, depending on whether the TF coherently activates (or represses) transcription, or whether it simultaneously influences transcription both positively and negatively. We show that responses to TFs acting in coherent mode are monotonic, and the overall effect can only change between activation and repression if there is a change in the mechanistic mode of TF action. This has been the

12

common understanding of duality in the literature, where the cellular context or the position of the TF binding site relative to other relevant sites determines its mechanism [1, 2, 3, 4].

In contrast, when TFs act incoherently, duality may be observed even if the TF acts always in the same (incoherent) mode. In this case, whether the TF acts as an activator or a repressor at the level of the steady-state mRNA response depends on the balance between the various processes involved in transcription. We distinguish three potential ways in which this balance can be tuned. First, the mRNA response can be tuned by changing the strength of the activating and repressive mechanistic effects of the TF, even if these changes are quantitative such that regulation remains incoherent. This would correspond to part of the observations of [30], where the position of the TF binding site with respect to the transcription start site determines the extent of the mechanistic effects of the TF. Second, the mRNA response may be tuned by factors other than the TF, while the mechanistic effect of the TF remains constant. In the cases exemplified in Fig. S2A and Fig S3A, the parameter that tunes the response can be interpreted to be related to the concentration of a chromatin regulator. Another relevant element could be the promoter strength and the associated dynamics of the transcriptional machinery, as recently studied theoretically by Ali et al. [31]. Finally, the third way to tune the mRNA response is by changing the occupancy of TF on DNA, as determined by the binding affinity or number of binding sites. This implies that the overall effect can change between activation and repression even if the molecular context in which the TF acts (other than its DNA binding sequence) is both qualitatively and quantitatively stable.

To illustrate our points, we have studied the regulated recruitment model (Fig. 1) and the regulated cycle model (Fig. 2). These two model architectures represent two fundamentally different model types. In the regulated recruitment model, the whole system can be assumed to operate at thermodynamic equilibrium (Fig. 1), as in the classic thermodynamic models of gene regulation, whereas in the regulated cycle model, irreversible transitions make the whole system non-equilibrium, although TF binding and polymerase recruitment can be assumed at equilibrium or not. This has allowed us to clarify the implications of a non-monotonic response. For a TF that binds to a single site, the response is always monotonic under assumptions of recruitment at thermodynamic equilibrium. In previous work by Gedeon et al. [35], it was shown that non-monotonicity could arise in equilibrium regulated recruitment models with more than one site, if the effects of the TF were site-dependent. More recently, it has been shown that non-monotonicity does not require more than one site. Instead, Mahdavi et al. [36] found non-monotonic responses for a single-site model of regulated recruitment when taken away from equilibrium (Fig. S1A). So, this could be taken as a suggestion that non-monotonicity may be a signature of underlying non-equilibrium regulated recruitment, a topic of active research [37, 38, 39, 40]. However, we have seen that this does not need to be the case. Equilibrium TF binding coupled to the regulation of the dissipative transcriptional cycle can already account for non-monotonic responses. So, these results suggest that non-monotonicity may be better regarded as an indication of incoherent regulation, although formally proving this for any general model remains open for future analyses.

In the light of this, we interpret the non-monotonicity and affinity-dependent response observed in our experimental data as evidence for incoherent regulation by SP1. We note that SP1 is mostly known for its transcriptional activating role, but repressive effects have also been described in the literature, even at the same gene [59, 60, 61].

Our work also provides a natural explanation for the "antagonism" between dual TF domains recently reported by Mukund et al. [29]. In that study, the authors examined the effect of chaining together TF domains. They found that when fusing two domains previously identified as dual (capable of activating transcription from a weak promoter, and repressing from a relatively stronger promoter), there was a tendency for "antagonism" at the weak promoter, with lower expression in the domain-domain constructs as compared to when each dual domain was chained to a neutral control sequence. The authors describe a series of experiments attempting to find explanations on the basis of lower TF expression, cellular stress etc, without success. Indeed, we suggest that the reason is likely that those "antagonistic" dual domains are acting incoherently, with simultaneous positive and negative effects on the transcription of the reporter, and that chaining together two domains is analogous to increasing binding site number in our non-monotonic or repressive regimes, where expression goes down with binding site number. Similarly, our work is consistent with the CRX regulation reported in [62], where activation or repression

depend on a combination of binding site number, affinity and promoter strength. In any case, there is always the possibility that the TF-interacting cofactors might change as a function of affinity, binding site number, or neighboring domains in the case of the domain-domain fusions, and careful molecular analysis should be performed to unequivocally pinpoint the mechanism. We hope our analyses will motivate such careful considerations in future studies.

It is tempting to speculate that the incoherent regulation studied here is a broader feature of TFs in natural systems. If this is the case, such that binding site number and affinity can tune the response direction, this might be another reason behind the widespread presence of low-affinity binding sites in eukaryotic genomes, a well-known fact that is often puzzling to interpret and has often been reasoned in terms of TF specificity [63, 64, 65, 66]. According to our findings, intermediate or low affinities might enable TFs to behave in the "right" direction, as required to functionally regulate their targets. If this is the case, this would also have evolutionary implications, as it would be easier to evolve the effect of a TF on a target gene by tuning binding site numbers or affinities, rather than evolving new networks of interacting coregulators. At the same time, tuning transcriptional response from activation to repression via changes in binding site number and affinity also poses questions regarding the robustness of the response to mutations. We hope that future work on incoherent TF regulation will clarify the molecular underpinnings of this phenomenon as well as its implications for our understanding of genomes and gene regulation.

4   Acknowledgements

5   Methods

5.1   Model simulations

The calculations of the fold change for a given parameter set and TF concentration x involve first calculating the steady state of the model for that parameter set and a TF concentration of 0, then the steady state at x, and then dividing the latter by the former. For equilibrium systems (Fig. 1A, S1B), the calculation of the steady-state for a given parameter set and TF concentration is done as follows [67, 37]. Choose a reference state (state 1) and assign $\mu_1 = 1$. Then for each state i, calculate μ_i as the product of edge labels (labels of equilibrium graphs, corresponding to edge label ratios of the original linear framework graph) from the reference state to i. The steady state probability of vertex i is then calculated as

$$P_i^* = \frac{\mu_i}{\Sigma_i \mu_i} \tag{7}$$

These calculations are done in Python.

For the non-equilibrium systems corresponding to the models in Fig. 1D and 2A, the steady state probability of each vertex i is calculated as

$$P_i^* = \frac{\rho_i}{\Sigma_i \rho_i} \tag{8}$$

where

$$\rho_i = \Sigma_{T \in \Theta_i} \Pi_T \tag{9}$$

where Θ_i is the set of spanning trees rooted at i, and Π_T is the product of the edge labels of tree T. The calculations are done in C++, using 100-digit precision floating-point types provided by the GNU MPFR Library through the Boost interface (`www.boost.org`). Further justification of the procedure and details for the models in the SI Appendix are given there.

The calculations of the hybrid system in Fig. 2D were done in Python. The following formula for the steady state mRNA was used (which can be obtained from the procedure just outlined):

$$m = \frac{k_3 k_1 k_2}{k_2 k_3 + k_3 k_{1r} + k_1 k_3 + k_1 k_2} \tag{10}$$

From this formula, the basal mRNA level was obtained with the basal rates, and the mRNA level at a given TF affinity, concentration and site number N was obtained by calculating a (average number of TF molecules bound), and subsequently the corresponding rate values, as detailed in Fig. 2D. In order to calculate a, we calculated the equilibrium steady-state probability of each binding configuration ν, P_ν^*, following Eq. 7. We then obtained a as $a = \sum_{i=0}^{N}(i \sum_{\#\nu=i} u_v^*)$, where $\#\nu$ means the number of sites bound in ν.

5.2 SynTF experiment

5.2.1 Cell culture and transfection

HEK293FT cells (Thermo Fisher Scientific) with stably integrated eGFP reporter were cultured as described in [56]. Transfection of synTF plasmid constructs was perfored using Lipofectamine 3000 (Thermo Fisher Scientific). 500,000 cells were plated in 6 cm culture plates and transfected the following day with the corresponding ng of synTF plus single stranded filler DNA (Thermo Fisher Scientific) to achieve equal amounts of transfected DNA of 1 μg DNA in total. After 24h cells were harvested and mRNA extracted using the RNeasy Mini Kit (Qiagen).

5.2.2 qRT-PCR

500 ng extracted total RNA was reverse transcribed into cDNA for each sample using Protoscript II reverse transcriptase (New England Biolabs) and oligo-dT primers (New England Biolabs). Quantitative real-time PCR was performed in triplicates using iTaqUniversal SYBRGreen reagent (Bio-Rad) on a CFX96 PCR machine (Bio-Rad). Primers were used in a final concentration of 243.2 nM. b-actin expression was used as a reference gene for relative quantification of RNA levels, as $2^{Cq_{actin} - Cq_{target}}$. For the GFP, fold change is expressed relative to a control condition where no plasmid is transfected (but there is still some GFP expression due to some basal promoter activity). For the ZF, fold change is expressed relative to the lowest input condition.

5.3 MPRA experiment

5.3.1 Sequence design

We generated cis-regulatory sequences (CRS) with different designs of SP1 binding sites, consisting of 1 to 6 binding sites, 3 orientations (only forward, only backward, tandem forward backward), 3 spacings (4, 10 and 20 base pairs) and 6 different affinities. To design binding sites of various affinities, we drew 100,000 random samples from the position weight matrix (PWM) of SP1_MOUSE.H11MO.1.A from hocomoco.v11 [68] and ranked the resulting sequences by the likelihood according to the original PWM. We selected 6 percentile values (10th, 25th, 50th, 75th, 90th and 100th quantile) and sampled sequences within $\pm 2.5\%$ of each value, except for the highest affinity, where sequences with the highest match were chosen. A total of 100 sequences for each design was generated computationally. TFBS were placed starting from the 3'end of the sequence, and spaces between the TFBS, as well as spaces between the last TFBS and the 5'end, were filled with random nucleotides, for a total length of 232 bases. 15 base pair adaptors were added 5'and 3'. The sequences were then screened for motif occurrences of other key hematopoietic transcription factors (the full list is provided in the SI) and sequences with strong binding sites for TFs in the background were excluded. In total 276 SP1 sequences were generated, containing all combinations of 1 to 6 binding sites, 3 orientations, 3 spacings and 6 different affinities. Additionally, 417 sequences containing only random DNA were synthesised.

5.3.2 lentiMPRA experiment

Oligos were synthesized at Twist Biosciences, for the experimental procedure the lentiMPRA protocol from [57] was followed. Lentivirus was produced in HEK293FT cells combining library plasmids with our cloned inserts (1.64 pM), psPAX2 (1.3 pM) and pMD2.G (0.72 pM) (Addgene: 137725, 12260 and 12259). 6h transfection was performed with Lipofectamine 3000 following manufactures protocol. Virus were collected 72h after transfection and precipitated with sucrose cushion ultracentrifugation (Boroujeni2018-xo). K562 cells were cultured in RPMI media supplemented with 10% FBS and 1% Penicillin-Streptomycin. 2 million K562 cells for both replicates were infected at a high MOI. Infection was stopped after 20 hours and cells were collected 3 days after infection. For nucleic acid isolation and library preparation we followed [57].

5.3.3 Data processing

GRE-Barcode association of the library and barcode counting at DNA and RNA level was performed using custom Perl scripts following [57]. The correlation of the replicates on RNA level and DNA level is 0.990 / 0.998. RNA counts per CRS were then normalized by DNA counts and we took the natural logarithm of this ratio as a final (raw) activity measure, with a correlation between replicates of 0.948. 197 sequences passed coverage filters. From the raw activity measure, the mean activity of sequences containing only random background DNA was subtracted, to achieve a scale were 0 is the activity induced by random DNA.

5.3.4 Sum of affinities score

We calculated for each site in a CRS the log-likelihood that it corresponds to an SP1 site, as described in Eq. S1 in [69], and summed over all the sites in the CRS. We considered the same PWM as during the design process, and a uniform background probability of 0.25 per nucleotide.

References

[1] C J Fry and P J Farnham. "Context-dependent transcriptional regulation". en. In: *J. Biol. Chem.* 274.42 (1999), pp. 29583–29586.

[2] David S Latchman. "Transcription factors: Bound to activate or repress". In: *Trends Biochem. Sci.* 26.4 (2001), pp. 211–213.

[3] Jun Ma. "Crossing the line between activation and repression". In: *Trends Genet.* 21.1 (2005), pp. 54–59.

[4] Patrick Boyle and Charles Després. "Dual-function transcription factors and their entourage: Unique and unifying themes governing two pathogenesis-related genes". In: *Plant Signal. Behav.* 5.6 (2010), pp. 629–634.

[5] D Pan and A J Courey. "The same dorsal binding site mediates both activation and repression in a context-dependent manner". In: *EMBO J.* 11.5 (1992), pp. 1837–1842.

[6] N Kirov et al. "Conversion of a silencer into an enhancer: evidence for a co-repressor in dorsal-mediated repression in Drosophila". In: *EMBO J.* 12.8 (1993), pp. 3193–3199.

[7] Todd Dubnicoff et al. "Conversion of Dorsal from an activator to a repressor by the global core-pressor Groucho". In: *Genes and Development* 11.22 (1997), pp. 2952–2957.

[8] Scott A Valentine et al. "Dorsal-Mediated Repression Requires the Formation of a Multiprotein Repression Complex at the Ventral Silencer". In: *Mol. Cell. Biol.* 18.11 (1998), pp. 6584–6594.

[9] Frank Sauer and Herbert Jäckle. "Concentration-dependent transcriptional activation or repression by Krüppel from a single binding site". In: *Nature* 353.6344 (1991), pp. 563–566.

[10] F Sauer and H Jäckle. "Dimerization and the control of transcription by Krüppel". en. In: *Nature* 364.6436 (1993), pp. 454–457.

[11] Frank Sauer et al. "Control of transcription by Krüppel through interactions with TFIIB and TFIIEβ". In: *Nature* 375.6527 (1995), pp. 162–164.

[12] Anna La Rosée-Borggreve et al. "hairy stripe 7 element mediates activation and repression in response to different domains and levels of Kruppel in the Drosophila embryo". In: *Mech. Dev.* 89.1-2 (1999), pp. 133–140.

[13] Emily R Weikum et al. "Glucocorticoid receptor control of transcription: precision and plasticity via allostery". In: *Nat. Rev. Mol. Cell Biol.* 18.3 (2017), pp. 159–174.

[14] Z Li et al. "Serine 574 phosphorylation alters transcriptional programming of FOXO3 by selectively enhancing apoptotic gene expression". In: *Cell Death Differ.* 23.4 (2016), pp. 583–595.

[15] Sarah Ross et al. "SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization". In: *Mol. Cell* 10.4 (2002), pp. 831–842.

[16] Corey Lourenco et al. "MYC protein interactors in gene transcription and cancer". en. In: *Nat. Rev. Cancer* 21.9 (2021), pp. 579–591.

[17] Thijs C J Verheul et al. "The Why of YY1: Mechanisms of Transcriptional Regulation by Yin Yang 1". In: *Frontiers in Cell and Developmental Biology* 8.September (2020), pp. 1–9.

[18] Haihong Zhong et al. "The phosphorylation status of nuclear NF-κB determines its association with CBP/p300 or HDAC-1". In: *Mol. Cell* 9.3 (2002), pp. 625–636.

[19] Vivien Ya Fan Wang et al. "The Transcriptional Specificity of NF-κB Dimers Is Coded within the κB DNA Response Elements". In: *Cell Rep.* 2.4 (2012), pp. 824–839.

[20] Timothy T Harden, Ben J Vincent, and Angela H DePace. "Transcriptional activators in the early Drosophila embryo perform different kinetic roles". en. In: *Cell Syst* 14.4 (2023), 258–272.e4.

[21] Dechen Fu and Jun Ma. "Interplay between positive and negative activities that influence the role of Bicoid in transcription". In: *Nucleic Acids Res.* 33.13 (2005), pp. 3985–3993.

[22] Chen Zhao et al. "The activity of the Drosophila morphogenetic protein Bicoid is inhibited by a domain located outside its homeodomain". In: *Development* 129.7 (2002), pp. 1669–1680.

[23] D Fu, C Zhao, and J Ma. "Enhancer Sequences Influence the Role of the Amino-Terminal Domain of Bicoid in Transcription". In: *Mol. Cell. Biol.* 23.13 (2003), pp. 4439–4448.

[24] Yael Katan, Reuven Agami, and Yosef Shaul. "The transcriptional activation and repression domains of RFX1, a context-dependent regulator, can mutually neutralize their activities". In: *Nucleic Acids Res.* 25.18 (1997), pp. 3621–3628.

[25] Etti Ben-Shushan et al. "Rex-1, a Gene Encoding a Transcription Factor Expressed in the Early Embryo, Is Regulated via Oct-3/4 and Oct-6 Binding to an Octamer Site and a Novel Protein, Rox-1, Binding to an Adjacent Site". In: *Mol. Cell. Biol.* 18.4 (1998), pp. 1866–1878.

[26] Josh Tycko et al. "High-Throughput Discovery and Characterization of Human Transcriptional Effectors". en. In: *Cell* 183.7 (2020), 2020–2035.e16.

[27] Andreas K Brödel et al. "Accelerated evolution of a minimal 63-amino acid dual transcription factor". en. In: *Sci Adv* 6.24 (2020), eaba2728.

[28] Nicole DelRosso et al. "Large-scale mapping and mutagenesis of human transcriptional effector domains". en. In: *Nature* 616.7956 (2023), pp. 365–372.

[29] Adi X Mukund et al. "High-throughput functional characterization of combinations of transcriptional activators and repressors". en. In: *Cell Syst* 14.9 (2023), 746–763.e5.

[30] Sunil Guharajan et al. "Quantifying the regulatory role of individual transcription factors in Escherichia coli". en. In: *Cell Rep.* 37.6 (2021), p. 109952.

[31] Md Zulfikar Ali et al. "Regulatory properties of transcription factors with diverse mechanistic function". en. In: *bioRxiv* (2023), p. 2023.06.15.545127.

[32] Mark Ptashne. "How eukaryotic transcriptional activators work". In: *Nature* 335.6192 (1988), pp. 683–689.

[33] S Natesan et al. "Transcriptional squelching re-examined". en. In: *Nature* 390.6658 (1997), pp. 349–350.

[34] Huey Lin et al. "Cellular toxicity induced by SRF-mediated transcriptional squelching". In: *Toxicol. Sci.* 96.1 (2007), pp. 83–91.

[35] Tomáš Gedeon et al. "When activators repress and repressors activate: A qualitative analysis of the Shea-Ackers model". In: *Bull. Math. Biol.* 70.6 (2008), pp. 1660–1683.

[36] Sara Mahdavi et al. "Flexibility and sensitivity in gene regulation out of equilibrium". en. In: *bioRxiv* (2023), p. 2023.04.11.536490.

[37] Javier Estrada et al. "Cellular Interrogation: Exploiting Cell-to-Cell Variability to Discriminate Regulatory Mechanisms in Oscillatory Signalling". In: *PLOS Computational Biology* 12.7 (2016), e1004995.

[38] Rok Grah, Benjamin Zoller, and Gašper Tkačik. "Nonequilibrium models of optimal enhancer function". In: *Proceedings of the National Academy of Sciences* 117.50 (2020), pp. 31614–31622.

[39] Jeehae Park et al. "Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity". In: *eLife* 8 (2019), e41266.

[40] Felix Wong and Jeremy Gunawardena. "Gene Regulation in and out of Equilibrium". In: *Annual Review of Biophysics* 49.1 (2020), pp. 199–226.

[41] Jeremy Gunawardena. "A Linear Framework for Time-Scale Separation in Nonlinear Biochemical Systems". In: *PLoS ONE* 7.5 (2012), e36321.

[42] Mark Ptashne. "Regulation of transcription: from lambda to eukaryotes". In: *Trends Biochem. Sci.* 30.6 (2005), pp. 275–279.

[43] T. Raveh-Sadka, Michal Levo, and Eran Segel. "Incorporating nucleosomes into thermodynamic models of transcription regulation". In: *Genome Research* 19.8 (2009), pp. 1480–1496.

[44] Leonid a Mirny. "Nucleosome-mediated cooperativity between transcription factors." In: *Proceedings of the National Academy of Sciences of the United States of America* 107.52 (2010), pp. 22534–22539.

[45] Javier Estrada et al. "Information Integration and Energy Expenditure in Gene Regulation". In: *Cell* 166.1 (2016), pp. 234–244.

[46] John W Biddle, Maximilian Nguyen, and Jeremy Gunawardena. "Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA". In: *eLife* 8 (2019), e41017.

[47] Gary K Ackers, A. D. Johnson, and Madeline A Shea. "Quantitative model for gene regulation by lambda phage repressor." In: *Proceedings of the National Academy of Sciences* 79.4 (1982), pp. 1129–1133.

[48] Lacramioara Bintu et al. "Transcriptional regulation by the numbers: models." In: *Current opinion in genetics & development* 15.2 (2005), pp. 116–24.

[49] Makiko Iwafuchi et al. "Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones". In: *Nat. Genet.* 52.4 (2020), pp. 418–427.

[50] Svetlana O Dodonova et al. "Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function". In: *Nature* 580.7805 (2020), pp. 669–672.

[51] Erin E. Swinstead et al. "Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically: A new perspective: Multiple transcription factors can effect chromatin pioneer functions through dynamic interactions with ATP-dependent chromatin remodeling factors". In: *BioEssays* 38.11 (2016), pp. 1150–1157.

[52] Julius Judd, Fabiana M Duarte, and John T Lis. "Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription". In: *Genes Dev.* 35.1-2 (2021), pp. 147–156.

[53] Isao Tamura et al. "Transcription factor C/EBPβ induces genome-wide H3K27ac and upregulates gene expression during decidualization of human endometrial stromal cells". In: *Mol. Cell. Endocrinol.* 520.July 2020 (2021), p. 111085.

[54] Yitzhak Reizel et al. "Collapse of the hepatic gene regulatory network in the absence of FoxA factors". In: *Genes and Development* 34.16 (2020), pp. 1039–1050.

[55] Iris Jonkers and John T Lis. "Getting up to speed with transcription elongation by RNA polymerase II". In: *Nat. Rev. Mol. Cell Biol.* 16.3 (2015), pp. 167–177.

[56] Rosa Martinez-Corral et al. "Transcriptional kinetic synergy: A complex landscape revealed by integrating modeling and synthetic biology". en. In: *Cell Syst* 14.4 (2023), 324–339.e7.

[57] M Grace Gordon et al. "lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements". en. In: *Nat. Protoc.* 15.8 (2020), pp. 2387–2412.

[58] Natalya Pya and Simon N Wood. "Shape constrained additive models". In: *Stat. Comput.* 25.3 (2015), pp. 543–559.

[59] Lanyue Hu et al. "Sp1 Mediates the Constitutive Expression and Repression of the PDSS2 Gene in Lung Cancer Cells". In: *Genes* 10.12 (2019), p. 977.

[60] Sylvain Fauquenoy et al. "Repression of Human T-lymphotropic virus type 1 Long Terminal Repeat sense transcription by Sp1 recruitment to novel Sp1 binding sites". en. In: *Sci. Rep.* 7 (2017), p. 43221.

[61] Faramarz Taheri and Haydee E P Bazan. "Platelet-activating factor overturns the transcriptional repressor disposition of Sp1 in the expression of MMP-9 in human corneal epithelial cells". en. In: *Invest. Ophthalmol. Vis. Sci.* 48.5 (2007), pp. 1931–1941.

[62] Michael A White et al. "A Simple Grammar Defines Activating and Repressing cis-Regulatory Elements in Photoreceptors". In: *Cell Rep.* 17.5 (2016), pp. 1247–1254.

[63] Andrea I Ramos and Scott Barolo. "Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution". In: *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368.1632 (2013).

[64] Emma K Farley et al. "Suboptimization of developmental enhancers". In: *Science* 350.6258 (2015), pp. 325–328.

[65] Justin Crocker et al. "Low affinity binding site clusters confer HOX specificity and regulatory robustness". In: *Cell* 160.1-2 (2015), pp. 191–203.

[66] Judith F Kribelbauer et al. "Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes". In: *Annu. Rev. Cell Dev. Biol.* 35.1 (2019), pp. 357–379.

[67] Inomzhon Mirzaev and Jeremy Gunawardena. "Laplacian Dynamics on General Graphs". In: *Bulletin of Mathematical Biology* 75.11 (2013), pp. 2118–2149.

[68] Ivan V Kulakovskiy et al. "HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis". en. In: *Nucleic Acids Res.* 46.D1 (2018), pp. D252–D259.

[69] Kenneth A Barr and John Reinitz. "A sequence level model of an intact locus predicts the location and function of nonadditive enhancers". In: *PLoS One* 12.7 (2017), p. 28715438.

Supplementary Information Appendix

1 Data and code availability

All the code to reproduce the plots of the paper is found in https://github.com/rosamc/TFduality.git.

2 Theoretical methods and calculations

2.1 Calculating steady states in the linear framework

The mathematical modelling of the paper follows the linear framework, which was introduced in [1] and subsequently developed [2] and applied to gene regulation for example in [3, 4]. These references, as well as the more recent review [5], can be consulted for further details and proofs, and here we only repeat the minimum required to understand the present work.

The linear framework assumes a timescale separation between slow and fast components of a biochemical a system. The fast components, in our case the various states of the regulatory region or bound polymerase, change over time. The slow components, in our case the TFs and RNA polymerase, are assumed to be in excess in a buffer, and their interaction (binding) with the fast components does not change their free concentration. In this setting, the fast components evolve over time following a finite-state, continuous-time, time-homogeneous Markov process, which has a corresponding linear framework graph: the vertices are the states of the system, the edges the transitions between them and the edge labels the infinitesimal transition rates, which can contain the slow components and have dimensions of $(\text{time})^{-1}$. The Master Equation for the time evolution of the probability of each state of the system is then given by:

$$\frac{d\vec{P}}{dt} = \mathcal{L}(G)\vec{P}, \tag{11}$$

where G is the graph, $\vec{P}$ is a column vector of state probabilities that sum to 1, and $\mathcal{L}(G)$ is the Laplacian matrix of G. The Laplacian matrix of G is a square matrix where (i, j), $i \neq j$ contains the label from j to i, and the diagonal terms contain the negative of the column sum. In the case of gene regulation, assuming ergodicity, P_i can be interpreted at the single cell and allele level as the average fraction of time the system spends in state i, whereas at the population level it can be interpreted as the fraction of cells in state i at a given time.

For strongly connected graphs as in the case of this paper (every vertex can be reached from any other vertex) the system always tends to a unique steady state $(d\vec{P}/dt = 0)$ up to a scalar multiple, which corresponds to the kernel of $\mathcal{L}(G)$. To calculate this steady state, we can exploit the graph, using the Matrix Tree Theorem.

A spanning tree of G is a subgraph that does not contain any cycle (tree) and spans all graph nodes (spanning). A spanning tree T is rooted at node i if it contains a path from any other node to node i. Let Π_T the product of the edge labels of a spanning tree T, Θ_i the set of all spanning trees rooted at i, and ν the set of all the vertices of the graph. A representative steady state $\rho(G)$, can be calculated using the

21

Matrix Tree Theorem, where each component of the steady state vector is given by:

$$\rho_i(G) = \Sigma_{T \in \Theta_i} \Pi_T \tag{12}$$

Because probabilities sum to 1, the steady-state probabilities of each state can be obtained by normalising the previous quantities:

$$P_i^*(G) = \frac{\rho_i}{\Sigma_{j \in \nu} \rho_j} \tag{13}$$

An example application of this formula to calculate the steady-state transcription rate is found in Fig. S1A. Using this formula, we computed the steady states of the non-equilibrium models of Fig. 1 and Fig. 2. Although it is in principle possible to also use it for 5-state cycle graph in Fig. S5 and the model in Fig. S8, the number of spanning trees becomes very large as the number of states in the cycle increases or with increasing number of sites. In those cases, for practical reasons we numerically solved for the kernel of the Laplacian using singular value decomposition. This was done with custom C++ code, using the SVD routine in Eigen 3.3.7 and 50-digit precision floating-point types provided by the GNU MPFR Library through the Boost interface (`www.boost.org`).

For equilibrium graphs, the calculations are substantially simpler, as explained next.

2.1.1 Equilibrium steady-states

When the biochemical system represented by a linear framework graph is assumed to be at thermodynamic equilibrium, the principle of detailed balance must be satisfied. This means that each pair of states, (i, j) is independently in flux balance: let P_i^* be the steady state probability of i, P_j^* the steady state probability of j, and $k_{i,j}$ the transition rate from i to j, then $P_i^* k_{i,j} = P_j^* k_{j,i}$. This implies, first, that all transitions must be reversible. Second, over any cycle, the product of the edge labels in the clockwise direction equals the product of the edge labels in the counter-clockwise direction (cycle condition). As a result, the quantities that specify the steady state are the ratios of forward and backward transition rates (which are related to the free energy difference between the two states), rather than the individual rates [4](Fig. S1). In other words, if we multiply by a factor ϵ a given rate, and multiply by the same factor ϵ its reverse rate, the steady state will remain unchanged.

Let $K_{m,n}$ be the ratio (label ratio) between the transition rates $k_{m,n}$ and $k_{n,m}$ between nodes m and n. Let ℓ_i be any path from a reference node to node i, and Π_{ℓ_i} the product of label ratios along this path. Then the steady state probability of node i is given by:

$$P_i^*(G) = \frac{\Pi_{l_i}}{\Sigma_{j \in \nu}(\Pi_{l_j})} \tag{14}$$

An example application of this formula to calculate the steady state transcription rate is found in Fig. S1B. We used this formula for the model in Fig. 1C and to calculate a in the model of Fig. 2D.

This is equivalent to the steady states obtained using the common statistical mechanics framework exploited by the thermodynamic models of gene regulation [6, 7]. This can be seen by considering the relationships between the label ratios $K_{m,n}$ and the free energy difference Δ_ϕ between m and n: $K_{m,n} = \exp\left(\frac{\Delta_\phi}{k_B T}\right)$, with k_B the Boltzmann's constant, and T the absolute temperature.

2.2 Monotonicity of the response for any recruitment model where a TF binds to a single site at thermodynamic equilibrium

Consider any model that assumes a given TF at concentration x regulates a gene by binding to a single site, the regulatory system is at thermodynamic equilibrium, and the steady-state transcription rate is taken to be a linear combination of the system steady-state probability distribution. No matter how complicated the system is, the response is always monotonic with respect to x.

To see this, consider an arbitrary system with some states with the TF bound and some states without the TF bound. A schematic of such a system is provided in Schema 1. Let Ψ the set of all nodes with TF bound, and Ω the set of all nodes without TF bound. As explained in the previous section, to calculate the steady state of each of the nodes i, we choose a path from a reference node to i. Let's consider the node in red as the reference. Then it is easy to see that any path from it to any node $i \in \Psi$ will be a product of label ratios that contains x one time i.e. will have a form of $\theta_i x$, where θ_i is a constant. Similarly, for any node $j \in \Omega$, the product of edge label ratios will be a constant θ_j (without x). Let $\psi \in \Psi$ the set of all the nodes with TF bound that contribute to expression, and $\omega \in \Omega$ the set of all the nodes without TF bound that contribute to expression. The transcription rate is of the form:

$$r(x) = \frac{A + Bx}{C + Dx} \tag{15}$$

with q_i the transcription rate from node i and $A = \sum_{i \in \omega} q_i \theta_i$, $B = \sum_{i \in \psi} q_i \theta_i$, $C = \sum_{i \in \Omega} \theta_i$, $D = \sum_{i \in \Psi} \theta_i$.

The derivative with respect to TF concentration x is then:

$$\frac{dr(x)}{dx} = \frac{B(C + Dx) - (A + Bx)D}{(C + Dx)^2} = \frac{BC + BDx - AD - BDx}{(C + Dx)^2} = \frac{BC - AD}{(C + Dx)^2} \tag{16}$$

Therefore, the sign of the derivative does not depend upon the TF concentration x, and therefore the response is only monotonically increasing or decreasing at equilibrium, depending on the sign of $BC - AD$. Whether the TF acts as an activator or repressor depends upon the various parameters of the system.
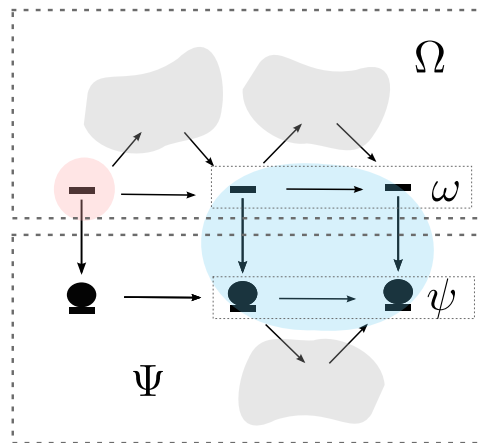
Schema 1. Cartoon that represents an arbitrary gene regulation model where a TF (disc) regulates a gene by binding at a single site, under thermodynamic equilibrium conditions. The states that contribute to expression are inside the blue region. The grey clouds denote arbitrary states and edges between them which do not include binding or unbinding of the TF. All transitions are reversible but edges are only shown in one direction, given that the relevant parameters to determine the steady-state behaviour is the ratio between forward and backward transitions. See text for details.

2.3 Monotonicity of the response for the two-conformation model at equilibrium (Fig. 1C)

The steady-state expression rate for the model in Fig. 1C is given by:

$$r^*(x) = q_3 P_3^*(x) + q_4 P_4^*(x) + q_7 P_7^* + q_8 P_8^* \tag{17}$$

We can easily show that the response to increasing TF concentration conforms to Eq. 15 and is therefore always monotonic at equilibrium. We consider node 1 in the graph of Fig. 1A as the reference node. This gives the following equation for $r^*(x)$:

$$r^*(x) = \frac{q_3 K_p + q_7 K_t \beta K_p + q_4 K_p K_T \omega x + q_8 K_t \alpha K_T \beta K_p \omega \gamma x}{1 + K_T x + K_p + K_T K_p \omega x + K_t + K_t \alpha K_T x + K_t \beta K_p + K_t \alpha K_T \beta K_p \omega \gamma x} \tag{18}$$

$$r^*(x) = \frac{A + Bx}{C + Dx} \tag{19}$$

with $\tag{20}$

$$A = q_3 K_p + q_7 K_t \beta K_p \tag{21}$$

$$B = q_4 K_p K_T \omega + q_8 K_t \alpha K_T \beta K_p \omega \gamma \tag{22}$$

$$C = K_p + K_t \beta K_p + 1 + K_t \tag{23}$$

$$D = K_p K_T \omega + K_t \alpha K_T \beta K_p \omega \gamma + K_T + K_t \alpha K_T \tag{24}$$

Next we discuss the parameteric conditions for monotonicity.

2.3.1 Direction of the response to a TF with coherent effects on polymerase for the two-conformation model at equilibrium

In the main text, we have assumed that the transcription rate is the same from all states with polymerase bound ($q_3 = q_4 = q_7 = q_8$). In this case, the biochemical activity of the TF is parameterized by α, ω and γ. If $\alpha > 1$, $\omega > 1$ and $\gamma \geq 1$, all the TF biochemical effects enhance polymerase binding, directly via the binding cooperativity ω, and indirectly through the effect on the closed/open balance through α ($\gamma \geq 1$ ensures that the cooperativity effect is in line with the opening effect). A more general case includes varying expression rates from each of the polymerase-bound states. In this case, an effect of the TF that aligns with the activating conditions for α, ω and γ defined above, requires that the transcription rate is at least as high from the polymerase-bound states in the open conformation as in the closed conformation, and that the binding of TF either doesn't affect transcription rate or enhances it (otherwise the effect would be inconsistent with the rest). This corresponds to the following ordering for the q_i: $q_3 \leq q_4 \leq q_7 \leq q_8$. In this case, the TF always behaves as an activator, as shown next.

The condition for the TF behaving as an activator (Eq. 16) is:

$$BC - AD > 0 \tag{25}$$

$$K_p K_T[\omega(q_4 + q_8 K_t \alpha \beta \gamma) K_p (1 + \beta K_t) + \omega(q_4 + q_8 K_t \alpha \beta \gamma)(1 + K_t) - \tag{26}$$
$$- (q_3 + q_7 K_t \beta)(1 + \alpha K_t) - (q_3 + q_7 K_t \beta) K_p \omega (1 + K_t \alpha \beta \gamma)] > 0$$

$$\omega(q_4 + q_8 K_t \alpha \beta \gamma)(1 + K_t) - (q_3 + q_7 K_t \beta)(1 + \alpha K_t) + \tag{27}$$
$$+ K_p \omega ((q_4 + q_8 K_t \alpha \beta \gamma)(1 + \beta K_t) - (q_3 + q_7 K_t \beta)(1 + K_t \alpha \beta \gamma)) > 0$$

$$q[\omega(\epsilon_4 + \epsilon_8 K_t \alpha \beta \gamma)(1 + K_t) - (1 + \epsilon_7 K_t \beta)(1 + \alpha K_t) + \tag{28}$$
$$+ K_p \omega ((\epsilon_4 + \epsilon_8 K_t \alpha \beta \gamma)(1 + \beta K_t) - (1 + \epsilon_7 K_t \beta)(1 + K_t \alpha \beta \gamma))] > 0$$

where we have eliminated a factor of $K_p K_T$ because it doesn't have an impact on the sign, reorganized and made the following substitutions: $q_3 = q$, $q_4 = \epsilon_4 q$, $q_7 = \epsilon_7 q$, $q_8 = \epsilon_8 q$. In order to prove that the TF acts as an activator if all its biochemical activities align, i.e. it promotes chromatin openness and polymerase binding ($\alpha > 1$, $\omega > 1$, $\gamma \geq 1$), and $\epsilon_i \geq 1$, $\epsilon_j \geq \epsilon_i$ if $j > i$, we will show that each of the lines in Eq. 28 is positive (remember that polymerase binding is favored in the open conformation: $\beta > 1$). We begin by the first line (we drop q since it does not impact the sign):

$$\omega(\epsilon_4 + \epsilon_8 K_t \alpha \beta \gamma)(1 + K_t) - (1 + \epsilon_7 K_t \beta)(1 + \alpha K_t) > 0 \tag{29}$$

$$\omega \epsilon_4 - 1 + \omega \epsilon_4 K_t - \alpha K_t + \omega \epsilon_8 K_t \alpha \beta \gamma + \omega \epsilon_8 K_t^2 \alpha \beta \gamma - \epsilon_7 K_t \beta - \epsilon_7 K_t^2 \alpha \beta > 0 \tag{30}$$

Given that $\omega > 1$, $\epsilon_4 \geq 1$, $\omega \epsilon_4 - 1 > 0$. For the rest of the terms in Eq. 30, we have

$$\omega \epsilon_4 K_t - \alpha K_t + \omega \epsilon_8 K_t \alpha \beta \gamma + \omega \epsilon_8 K_t^2 \alpha \beta \gamma - \epsilon_7 K_t \beta - \epsilon_7 K_t^2 \alpha \beta > 0 \tag{31}$$

$$\omega \epsilon_8 K_t \alpha \beta \gamma - \epsilon_7 K_t \alpha \beta + \omega \epsilon_4 + \omega \gamma \overline{\epsilon_8}(\overline{\beta} \alpha) - \alpha - \overline{\beta} > 0 \tag{32}$$

$$[\alpha \beta (\omega \epsilon_8 K_t \gamma - \epsilon_7 K_t)] + [\omega \epsilon_4 + \gamma \overline{\epsilon_8}(\overline{\beta} \alpha) - (\alpha + \overline{\beta})] > 0 \tag{33}$$

with $\epsilon_8 = \overline{\epsilon_8} \epsilon_7 \geq 1$ and $\overline{\beta} = \epsilon_7 \beta > 1$. It is easy to see that the terms inside the first square brackets result in a positive term. To see that the same happens for the second square bracket, note that $1 + \overline{\beta} \alpha > \alpha + \overline{\beta}$ if $\alpha > 1$ and $\overline{\beta} > 1$. To see this, let $\alpha = 1 + \theta_a$ and $\overline{\beta} = 1 + \theta_b$, with $\theta_a > 0$, $\theta_b > 0$, then:

$$1 + (1 + \theta_a)(1 + \theta_b) > (1 + \theta_a) + (1 + \theta_b) \tag{34}$$

$$1 + 1 + \theta_a + \theta_b + \theta_a \theta_b > 1 + \theta_a + 1 + \theta_b \tag{35}$$

$$\theta_a \theta_b > 0 \tag{36}$$

With a similar reasoning we can show that the second line of Eq. 28 is positive under the parametric assumptions considered:

$$\epsilon_4 - 1 + \epsilon_4 \beta K_t + \epsilon_8 K_t \alpha \beta \gamma - K_t \alpha \beta \gamma - \epsilon_7 K_t \beta + \epsilon_8 K_t^2 \beta^2 \alpha \gamma - \epsilon_7 K_t^2 \beta^2 \alpha \gamma > 0 \tag{37}$$

Clearly $\epsilon_4 - 1 \geq 1$ and $\epsilon_8 K_t^2 \beta^2 \alpha \gamma \geq \epsilon_7 K_t^2 \beta^2 \alpha \gamma$ since $\epsilon_8 \geq \epsilon_7$. For the rest, dropping one factor of $K_t \beta$ per term:

$$\epsilon_4 + \epsilon_8(\alpha \gamma) - (\epsilon_7 + \alpha \gamma) > 0 \tag{38}$$

Since $\epsilon_8 \geq \epsilon_7$ and $\epsilon_4 \geq 1$, positivity follows seeing that the expression is at least of the form $1 + ab > a + b$ with $a \equiv \epsilon_8$ and $b \equiv \alpha \gamma$.

If the TF is no longer assumed to always enhance polymerase binding, for example if $\omega < 1$, the cooperativities misalign between conformations, or the relationships between the q_i are such that expression rate does not follow chromatin openness and TF binding, then the above inequalities will not always hold. Instead, relationships between the various parameters determine whether increases in TF concentration

result in more or less transcription, as shown in Fig. 1D and Fig. S2.

On the other hand, repression is ensured (in this case $BC - AD < 0$) if, as before, polymerase has higher affinity for the open conformation ($\beta > 1$), but the TF hampers polymerase binding so that $0 < \alpha < 1, 0 < \omega < 1, 0 < \gamma < 1$, and expression is either equal from all polymerase bound states ($\epsilon_4 = 1, \epsilon_7 = 1, \epsilon_8 = 1$), or higher from the open conformation but reduced in the presence of TF ($0 < \epsilon_4 < 1, 1 < \epsilon_8 < \epsilon_7$).

Let's start with the case $\epsilon_i = 1$. In this case, the second line in Eq. 28 vanishes. To see that the first line is negative, notice that in Eq. 30, $\omega \epsilon_4 - 1 < 0$ if $0 < \omega < 1$ and $\epsilon_4 = 1$. For the rest, let's focus on the left-hand side of Eq. 33. Notice that $1 + ab < a + b$ if $0 < a < 1, b > 1$, which can be easily shown if $a \equiv 1 - \theta_a, 0 < \theta_a < 1, b \equiv 1 + \theta_b, \theta_b > 0$:

$$1 + (1 - \theta_a)(1 + \theta_b) < (1 - \theta_a) + (1 + \theta_b) \tag{39}$$
$$1 + 1 - \theta_a + \theta_b - \theta_a \theta_b < 1 - \theta_a + 1 + \theta_b \tag{40}$$
$$2 - \theta_a + \theta_b - \theta_a \theta_b < 2 - \theta_a + \theta_b \tag{41}$$

It is then easy to see that the term in the second square brackets in Eq. 33 must be negative. For the other terms, it is easy to see that they are also negative.

When the ϵ_i are no longer 1 but satisfy the constraints mentioned ($0 < \epsilon_4 < 1, 1 < \epsilon_8 < \epsilon_7$) (now $0 < \overline{\epsilon_8} < 1$) similar reasonings allow to show that both lines in the left-hand side of Eq. 28 are negative.

2.3.2 Direction of the response if the TF that acts on α or ω only

If all the q_i are equal, $\gamma = 1$, and the TF only acts through α or ω, then it is easy to see that only activation or repression arise depending on whether the parameter is greater or less than 1. In this case, Eq. 28 simplifies to:

$$q[\omega(1 + K_t \alpha \beta)(1 + K_t) - (1 + K_t \beta)(1 + \alpha K_t)] > 0 \tag{42}$$
$$\omega + \omega K_t + K_t \alpha \beta \omega + K_t^2 \alpha \beta \omega - 1 - \alpha K_t - \beta K_t - \alpha \beta K_t^2 > 0 \tag{43}$$
$$(\omega - 1) + \alpha \beta K_t^2 (\omega - 1) + K_t(\omega + \omega \alpha \beta - (\alpha + \beta)) > 0 \tag{44}$$

If $\omega = 1$ (TF acts through α), the condition further reduces to

$$1 + \alpha \beta - (\alpha + \beta) > 0 \tag{45}$$

Remember we consider always $\beta > 1$. If $\alpha > 1$, the condition is fulfilled, which means the TF behaves as an activator. To see this, let $\alpha \equiv 1 + \theta_a$, with $\theta_a > 0$:

$$1 + (1 + \theta_a)\beta > (1 + \theta_a + \beta) \tag{46}$$
$$1 + \beta \theta_a + \beta > 1 + \theta_a + \beta \tag{47}$$

On the contrary, if $0 < \alpha < 1$ ($\theta_a < 0$) the left-hand side of Eq. 45 is negative, and the TF behaves as a repressor.

Similarly, if $\alpha = 1$, the condition in Eq. 44 becomes:

$$(\omega - 1) + \beta K_t^2 (\omega - 1) + K_t((\omega - 1) + \beta(\omega - 1)) > 0 \tag{48}$$

Clearly, the condition is fulfilled if $\omega > 1$, whereas the left-hand side becomes negative if $\omega < 1$.

2.4 Simplifying assumptions of the TF-chromatin model away from equilibrium (Fig. 1E)

For simplicity, we consider in the main text the extreme where the TF binds with the same kinetics to the two conformations, and can modulate the chromatin opening rate by a factor α. This could be interpreted as the nucleosome occupying the transcription start site but not the TF binding site. In addition, we assume the TF and polymerase can affect each other's binding, maintaining the cycle condition in the cycles encompassing only the binding and unbinding transitions ($\{1, 2, 3, 4\}$ and $\{5, 6, 7, 8\}$). This is represented by a cooperativity parameter ω, which is modelled as modifying the unbinding rate of the TF and polymerase from the states 4 and 8, with both bound. γ accounts for a potential difference in the binding cooperativity between the two conformations. In the main text, we consider $\gamma = 1$. Positive binding cooperativity, which favors binding, is given by $\omega > 1$, and negative binding cooperativity is given by $0 < \omega < 1$. The overall transcription rate is assumed to be given by the weighted average of the transcription rate from each state with polymerase bound, and in the main text we assume that it is equal across states ($q_3 = q_4 = q_7 = q_8 = q$).

2.5 Monotonicity of the response for the TF-chromatin model away from equilibrium (Fig. 1E)

Considering the previously-stated constraints on the parameter relationships, here we present the reasoning for analytically proving the monotonicity conditions. For the general case without constraints (Fig. S4) we have only performed a numerical exploration, as explained in the next subsection.

2.5.1 Activation

It can be proved for the non-equilibrium model in Fig. 1E that if $\alpha > 1$, $0 < \beta < 1$, $\omega > 1$, $\gamma \geq 1$, and $q_3 = q_4 = q_7 = q_8 = 1$, then

$$\frac{d(P_3(x) + P_4(x) + P_7(x) + P_8(x))}{dx} > 0 \tag{49}$$

which means that the response to the TF is monotonically increasing when polymerase has higher affinity for the open conformation ($\beta < 1$), the TF enhances polymerase binding both by $\alpha > 1$ and $\omega > 1$, $\gamma \geq 1$, and expression is the same from all polymerase-bound states.

In this case, there are a large number of terms in the corresponding gene regulatory function and derivative. In order to show monotonicity, we use the following observation:

Consider two polynomials in x of degree n, and their corresponding derivatives ($'$ denotes derivative with respect to x):

$$A = \sum_{k=0}^{n} a_k x^k \rightarrow A' = \sum_{k=1}^{n} a_k \, k \, x^{k-1} \tag{50}$$

$$B = \sum_{k=0}^{n} b_k x^k \rightarrow B' = \sum_{k=1}^{n} b_k \, k \, x^{k-1} \tag{51}$$

$$\tag{52}$$

27

Now consider the derivative with respect to x of the ratio A/B:

$$\left(\frac{A}{B}\right)' = \frac{1}{B^2}(A'\,B - A\,B') \tag{53}$$

The sign is determined by the term in the parenthesis:

$$P = A'\,B - A\,B' = \sum_{k=1}^{n} a_k\,k\,x^{k-1} \sum_{k=0}^{n} b_k x^k - \sum_{k=0}^{n} a_k x^k \sum_{k=1}^{n} b_k\,k\,x^{k-1} \tag{54}$$

Expanding it and rearranging, P can be expressed as:

$$P = 0\,(...) - a_0(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1}) +$$
$$a_1\,(b_0 + b_1\,x + b_2\,x^2 + ... + b_{n-1}x^{n-1} + b_n\,x^n) - a_1\,x(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1}) +$$
$$2\,a_2\,x\,(b_0 + b_1\,x + b_2\,x^2 + ... + b_{n-1}x^{n-1} + b_n\,x^n) - a_2 x^2\,(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1}) +$$
$$3\,a_3\,x^2\,(b_0 + b_1\,x + b_2\,x^2 + ... + b_{n-1}x^{n-1} + b_n\,x^n) - a_3 x^3\,(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1}) +$$
$$\vdots$$
$$(n-1)\,a_{n-1}\,x^{n-2}\,(b_0 + b_1\,x + b_2\,x^2 + ... + b_{n-1}x^{n-1} + b_n\,x^n) - a_{n-1}x^{n-1}(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1}) +$$
$$n\,a_n\,x^{n-1}\,(b_0 + b_1\,x + b_2\,x^2 + ... + b_{n-1}x^{n-1} + b_n\,x^n) - a_n x^n(b_1 + 2\,b_2 x + 3\,b_3 x^2 + ... + n\,b_n x^{n-1})$$

Terms can be collected according to the subindices of the coefficients a and b and the exponent of x. For example:

- Subindices 0 and 1, exponent 0: $b_0\,a_1 - a_0\,b_1$

- 1 and 2, exponent 2: $a_1\,b_2\,x^2 - a_1\,x\,2\,b_2\,x + 2\,a_2\,x\,b_1\,x - a_2\,x^2\,b_1 = (b_1\,a_2 - a_1\,b_2)x^2$

- 2 and n-1, exponent n: $2\,a_2\,x\,b_{n-1}x^{n-1} - a_2\,x^2(n-1)b_{n-1}x^{n-2} + (n-1)a_{n-1}x^{n-2}b_2\,x^2 - 2a_{n-1}x^{n-1}b_2\,x = (n-1-2)(b_2\,a_{n-1} - a_2\,b_{n-1})x^n$

We notice that the terms where the subindices are the same cancel, and so the above grouping leads to the following expression for P:

$$P = \sum_{i=0...n-1,j=1...n,i<j} (j-i)(b_i\,a_j - a_i\,b_j)x^{i+j-1} \tag{55}$$

Thus, the derivative is **positive** if

$$b_i\,a_j - a_i\,b_j > 0 \quad \forall\{i,j\},(i<j) \tag{56}$$

In this case, or if some terms vanish and others fulfill the condition, monotonicity is ensured.

We have implemented an algorithm that (for the graph in Fig. 1E):

1) Computes the spanning trees rooted at each node and from those, the polynomials for the numerator and denominator of the steady-state response function $r^*(x)$.

2) For each pair of i, j, $i = 0, ..., 3$, $j = 1, ..., 4$ computes the corresponding $c = b_i\,a_j - a_i\,b_j$.

3) Splits each c into expressions with a common factor in $a_T, b_T, a_p, b_p, k_o, k_c$.

This results into a total of 2673 expressions in $\alpha, \beta, \gamma, \omega$. We use mathematica to check that each of them is

individually positive for the parameter constraints described above, and therefore the whole derivative is positive. Note that in the code for the proofs, ω and γ represent $1/\omega$ and $1/\gamma$ in the paper. Therefore the mathematica code requires $\omega < 1$, $\gamma \leq 1$ to represent activating conditions.

Using the same approach, we can show that

$$\frac{d(P_8(x))}{dx} > 0 \tag{57}$$

$$\frac{d(P_7(x) + P_8(x))}{dx} > 0 \tag{58}$$

$$\frac{d(P_4(x) + P_7(x) + P_8(x))}{dx} > 0 \tag{59}$$

for the same parametric constraints for α, β, ω, γ.

Notably, Eq. 49 and Eq. 57- 59 imply that the response is also monotonic for the more general case where expression is not the same from all states, but increases with the TF bound and the open conformation such that $q_3 \leq q_4 \leq q_7 \leq q_8$.

To see this, first note that showing that

$$\frac{d(q_3 P_3(x) + q_4 P_4(x) + q_7 P_7(x) + q_8 P_8(x))}{dx} > 0 \tag{60}$$

if $q_3 \leq q_4 \leq q_7 \leq q_8$ is equivalent to showing that

$$\frac{d(q_3 P_3(x) + q_3 \epsilon_1 P_4(x) + q_3 \epsilon_1 \epsilon_2 P_7(x) + q_3 \epsilon_1 \epsilon_2 \epsilon_3 P_8(x))}{dx} > 0 \tag{61}$$

for $\epsilon_i \geq 1 \forall i$, $q_3 > 0$. In turn, this is equivalent to showing that

$$\frac{d(P_3(x) + \epsilon_1 P_4(x) + \epsilon_1 \epsilon_2 P_7(x) + \epsilon_1 \epsilon_2 \epsilon_3 P_8(x))}{dx} > 0 \tag{62}$$

for $\epsilon_i \geq 1 \forall i$.

Let $P_i' \equiv dP_i(x)/dx$. From Eq. 49:

$$P_3' + P_4' + P_7' + P_8' > 0 \tag{63}$$

$$P_4' + P_7' + P_8' > -P_3' \tag{64}$$

$$\epsilon_1 (P_4' + P_7' + P_8') > -P_3' \tag{65}$$

Where the last inequality must hold given $\epsilon_1 \geq 1$ and Eq. 59. Moreover, given Eq. 58

$$\epsilon_1 (P_7' + P_8') > -\epsilon_1 P_4' - P_3' \tag{66}$$

$$\epsilon_2 (\epsilon_1 (P_7' + P_8')) > -\epsilon_1 P_4' - P_3' \tag{67}$$

$$P_3' + \epsilon_1 P_4' + \epsilon_1 \epsilon_2 P_7' + \epsilon_1 \epsilon_2 P_8' > 0 \tag{68}$$

Finally, since we also know $P_8' > 0$, Eq. 62 must hold for $\epsilon_3 \geq 1$.

2.5.2 Repression

The same reasoning and algorithm can be used to proof the monotonically decreasing conditions.

A repressive TF can be encoded in the model as reducing the opening rate ($\alpha < 1$) and having negative cooperativity with polymerase ($\omega < 1$), with $\gamma = 1$. Under these conditions, it can be shown that:

$$\frac{d(P_7(x))}{dx} < 0 \tag{69}$$

$$\frac{d(P_7(x) + P_8(x))}{dx} < 0 \tag{70}$$

$$\frac{d(P_3(x) + P_7(x) + P_8(x))}{dx} < 0 \tag{71}$$

$$\frac{d(P_4(x) + P_3(x) + P_7(x) + P_8(x))}{dx} < 0 \tag{72}$$

If the repressive role of the TF is also manifested in the expression rates, such that $q_4 \leq q_3 \leq q_8 \leq q_7$ the response continues to be monotonically decreasing, that is

$$q_4 P_4' + q_4 \epsilon_1 P_3' + q_4 \epsilon_1 \epsilon_2 P_8' + q_4 \epsilon_1 \epsilon_2 \epsilon_3 P_7' < 0 \tag{73}$$

with $\epsilon_i \geq 1$.

To verify this, note that we can do the same reasoning as above given Eqs. 69-72:

$$P_4' + P_3' + P_8' + P_7' < 0 \tag{74}$$

$$P_3' + P_8' + P_7' < -P_4' \tag{75}$$

$$\epsilon_1(P_3' + P_8' + P_7') < -P_4' \tag{76}$$

$$\epsilon_2(\epsilon_1(P_8' + P_7')) < -P_4' - \epsilon_1 P_3' \tag{77}$$

$$\epsilon_1 \epsilon_2 \epsilon_3 P_7' < -P_4' - \epsilon_1 P_3' - \epsilon_1 \epsilon_2 P_8' \tag{78}$$

We note that when $q_3 = q_4 = q_7 = q_8$ and $\gamma \neq 1$, or $\gamma = 1$ but the expression rates are not equal nor satisfy the ordering considered in the above discussion (or $\gamma \neq 1$ and the expression rates are not equal nor satisfy the orderings) then activation or repression is not ensured only on the basis of α and ω, and non-monotonic responses can appear even if both of these parameters align (Fig. S3B-E).

2.5.3   A more general non-equilibrium TF-chromatin model

In the main text and in the previous sections, we discussed the TF-chromatin model under certain constraints imposing the cycle condition in the polymerase effects on chromatin, and cooperativity being dictated only by a change in the unbinding rate. We can also investigate a less constrained model where binding on-rates vary between conformations, the cycle condition does not hold in any of the cycles, and both the TF and polymerase can affect both the opening and closing rates by arbitrary factors (Fig. S4).

In this case, we have not been able to achieve an analytical proof as above, so we resorted to extensive numerical exploration for random parameter sets to determine the monotonicity constraints in Fig. S4. When these constraints are not satisfied, non-monotonic responses appear (Fig. S4B, "incoherent mode").

For monotonic activation, the constraints imply that 1) both the TF and polymerase exhibit positive cooperativity (higher binding rate and lower unbinding rate when binding to or from the state where the other is bound, as compared to binding and unbinding when the other is not bound in the same conformation); 2) both the TF and polymerase bind more in the open conformation (on-rates higher and off-rates lower for the states in the open conformation than for any of the closed); 3) expression is higher from the states with both TF and polymerase and even higher in the open conformation; 4) both TF and polymerase favor the open conformation.

For monotonic repression, polymerase binds more in the open conformation and favors it, and expresses

more from the open conformation. The repressive effect of the TF comes from: 1) TF binds more in the closed conformation and it tends to favor the closed conformation; 2) TF reduces polymerase binding (and vice-versa, i.e. there is negative binding cooperativity between the two); 3) when both TF and polymerase are bound, the effect of the TF on the conformational transitions is stronger than that of polymerase, so closing is favored.

To test that these constraints ensure monotonic responses, we sampled 1 million parameter sets for each of the two conditions. The parameter sets were sampled uniformly on a base 10 log-scale, from the polytopes defined by the constraints, within the bounds $(0,6)$. For this, we used the code in https://github.com/kmnam/convex-polytopes.git (commit 479f2e4e675803e213d44e6361f06491f336b0b7). Then, parameters were obtained within the range (1e-3, 1e3) after rescaling and exponentiating using base 10.

For each parameter set, monotonicity was assessed by numerically searching for zeros of the derivative of the input-output function $r^*(x)$. This was done with custom C++ code with high precision types provided by Boost/MPFR (100 digit precision was used). This code is run from Python using pybind11, and is based on the class GRF-Calculations available in https://github.com/rosamc/GeneRegulatoryFunctions.git (commit 5822641c807b952ea9be3cc4e960a9ba134bcfd7a). First, $r^*(x)$ was defined as a rational function in terms of the q_i and the graph edge labels (as described in "Calculating steady states in the linear framework"). Then, for each parameter set, the numerical values of the coefficients of the rational function were obtained, and the derivative $d(r^*(x))/dx$ calculated using the derivative formula. We searched for zeros of this derivative using the custom implementation of the Aberth-Ehrlich polynomial root-finding method available in https://github.com/kmnam/polynomials.git (commit af5a8318a6d637680033858a9b902c6d02eff613).

2.6 Monotonicity of the response in the transcription cycle model for TFs that only enhance or hamper the cycle

A similar procedure to that described for the non-equilibrium TF-chromatin model can be applied to show that the response to the TF in the model of Fig. 2C is monotonically increasing if $\epsilon_1 \geq 1$, $\epsilon_2 \geq 1$, $\epsilon_3 \geq 1$, $\epsilon_{1_r} \leq 1$ with at least one inequality. In this case, we work directly with the numerator of the derivative of the GRF. We collect terms with common binding, unbinding and basal rate parameters. This leaves a collection of expressions in ϵ_i, that have to be shown always positive when the abovementioned constraints hold. We use mathematica to do so. When the inequalities are reversed, we find that monotonic repression is ensured.

2.7 Variations of the transcription cycle model

In order to check that the emergence of non-monotonicity in the model in Fig. 2C generalises beyond the specific details of that particular case, we explored variations of the model considering different order of the processes controlled positively and negatively by the TF, reversibility patterns, points of control of the TF, and number of states (in addition to the model with 3 states, we tested a model with 5 states). In order to refer to each model, we use a string, with general form $S\_R\_t = T\_d = d_1, d_2$ (Fig. S5A), which corresponds to the following encoding:

S is the number of states in the cycle of the model (we tested $S \in \{3, 5\}$). We consider a labelling of the states $1, 2, \cdots, S-1, S$ in a clockwise manner, such that there is an edge between i and $i+1$ (or i and 1 if $i = S$), which we refer to as transition i. As in the model in Fig. 2C, transcription rate is proportional to the flux through transition S, added for both TF bound and unbound states. To refer to a transition in the forward direction, we use subscript $_f$, and if the transition is reversible, we use subscript $_r$ to refer to the reverse transition. R is the set of reversible transitions in the cycle ($R \subseteq \{1, 2, \cdots ..., S-1\}$), and

T the set of two transitions assumed to be modulated by the TF ($T = \{i_x, j_y\}$, $x, y \in \{_f, _r\}$). Finally we use "d_1, d_2" to encode the direction in which the TF affects the first transition and the second transition in T, such that $d_i = p$ if the value of that transition rate is higher when the TF is bound as compared to the basal cycle, and $d_i = m$ if the value of the transition rate is lower when the TF is bound. The model in Fig. 2C, with the parameters in that panel, corresponds to $3\_1\_t = \{1_f, 2_f\}\_d = p, m$ (Fig. S5A).

We explored the following models:

1. $3\_\{1\}\_t = \{1_f, 2_f\}\_d = m, p$
2. $3\_\{1\}\_t = \{1_r, 2_f\}\_d = p, p$
3. $3\_\{1\}\_t = \{1_r, 2_f\}\_d = m, m$
4. $3\_\{1\}\_t = \{1_r, 3_f\}\_d = p, p$
5. $3\_\{1\}\_t = \{1_r, 3_f\}\_d = m, m$
6. $3\_\{2\}\_t = \{1_f, 2_f\}\_d = p, m$
7. $3\_\{2\}\_t = \{1_f, 2_f\}\_d = m, p$
8. $3\_\{2\}\_t = \{1_f, 2_r\}\_d = p, p$
9. $3\_\{2\}\_t = \{1_f, 2_r\}\_d = m, m$
10. $3\_\{2\}\_t = \{1_f, 3_f\}\_d = m, p$
11. $3\_\{2\}\_t = \{1_f, 3_f\}\_d = p, m$
12. $3\_\{1, 2\}\_t = \{1_f, 2_f\}\_d = p, m$
13. $3\_\{1, 2\}\_t = \{1_f, 2_f\}\_d = m, p$
14. $3\_\{1, 2\}\_t = \{1_f, 3_f\}\_d = p, m$
15. $3\_\{1, 2\}\_t = \{1_f, 3_f\}\_d = m, p$
16. $5\_\{1, 2, 3, 4\}\_t = \{1_f, 2_f\}\_d = p, m$
17. $5\_\{1, 2, 3, 4\}_t = \{2_f, 4_f\}\_d = m, p$

For each model, we explored the parameter space aiming to identify particular kinds of behaviour: monotonically increasing and decreasing responses, bell-shaped responses, u-shaped responses and responses with more than one critical point. For this, we developed a score for input-output responses such that each of these kinds of responses falls into a particular region of a 2D space:

Given the steady-state response to TF $r^*(x)$, we evaluate it over the relevant range $[x_0, x_1]$, and consider $f(x) = log_2(r^*(x)/r^*(0))$. Then we define y_1 and y_c as follows:

$$y_1 \equiv f(x_1)$$

$$y_c \equiv \begin{cases} y_1 & \text{if } f(x) \text{ is monotonically increasing} \\ f(x_0) & \text{if } f(x) \text{ is monotonically decreasing} \\ f(x_c) & \text{if } f(x) \text{ has a single critical point at } x = x_c \\ 0 & \text{if } f(x) \text{ has more than one critical point} \end{cases}$$

Note that $f(x_0) = 0$ always. Therefore, a given response corresponds to a point in a 2D space defined by (y_c, y_1), as follows:

32

- on the positive y-axis: more than one critical point.
- diagonal of the positive quadrant: monotonically increasing.
- below the diagonal of the positive quadrant: bell-shaped ending at $f(x) > 0$.
- quadrant where $y_c > 0, y_1 < 0$: bell-shaped ending at $f(x) < 0$.
- on the negative y-axis: monotonically decreasing.
- quadrant where $y_c < 0, y_1 < 0$: u-shaped, ending at $f(x) < 0$.
- quadrant where $y_c < 0, y_1 > 0$: u-shaped, ending at $f(x) > 0$.

We then used a biased-sampling algorithm that we've developed in previous work [8] to explore the region of this 2D space that each model can occupy. Briefly, the 2D space is discretised into a grid, and the algorithm samples parameter sets iteratively, calculates its (y_c, y_1) value and corresponding grid cell, and iteratively modifies the parameter sets to fill more grids, until no new grid cells can be filled. We note that for $S = 5$ the exploration is very slow, so we stopped after 30 days of exploration. Parameter values for the basal cycle transitions and binding and unbinding rates were sampled within the range $[1, 10^4]$. For the TF-modulated transitions, the TF was assumed to modulate the basal rate a hundred fold (up or down).

We observed the emergence of non-monotonic, bell-shaped responses, in all models tested, except when the two transitions regulated by the TF start at the same cycle state (models $3\_\{1\}\_t = \{1_r, 2_f\}\_d = p, p$, $3\_\{1\}\_t = \{1_r, 2_f\}\_d = m, m$). Moreover, for model $5\_\{1, 2, 3, 4\}\_t = \{2_f, 4_f\}\_d = m, p$, we also found non-monotonic u-shaped responses. Examples of non-monotonic responses for various models are shown in Fig. S5B-E, showing how very similar bell-shaped responses can be obtained by various implementations of the model.

2.8   Incoherent regulation through functional interference among activators

The incoherent regulatory mode of a TF, in which it acts positively on some transcriptional step(s) and negatively on other(s), could arise through various molecular mechanisms. For example, we could imagine that negative effects can arise for a TF that on its own is always activating, as a result of interference with another activator. To illustrate this scenario, we expand the polymerase cycle model in Fig. 2C to include a second TF, Y (Fig. S8). We assume that X and Y bind each to their own site, independently of each other, and each has a given on- and off- binding rate, the same for all the cycle states. When bound individually, TF X is assumed to enhance the first and second transitions of the cycle, from basal values $(k_{1,\emptyset}, k_{2,\emptyset})$, to values $(k_{1,X}, k_{2,X})$, and TF Y is assumed to enhance the second transition only. Therefore, when present alone, they always behave as activators. However, we could imagine a situation where they interfere with each other's effect on the second transition when bound together, i.e. $k_{2,\emptyset} < k_{2,\{X,Y\}} < \min(k_{2,X}, k_{2,Y})$. In this situation, Fig. S8B shows that X always behaves as an activator when bound alone (gray line) but in the presence of a sufficient concentration of Y (black line), X can behave as an activator, a repressor or cause a non-monotonic response over a fixed concentration range, with the X unbinding rate tuning the direction of the response. We see that this is the same behaviour that we've seen for the model in Fig. 2C with a single TF considered explicitly, when it is assumed to act in incoherent mode, suggesting that functional interference might be a plausible molecular mechanism underlying incoherent regulation.

3   Supplemental Experimental Materials and Methods

3.1   eGFP reporter construct for synTF experiments

The reporter construct is the same as that in [9]. It consists of a single synthetic zinc finger binding site (CGGCGTAGCCGATGTCGCGC) upstream of a minimal CMV promoter (taggcgtgtacggtgggaggc-

ctatataagcagagctcgtttagtgaaccgtcagatcgcctgga) driving d2EGFP (EGFP destabilized with signal peptide for fast degradation (fusion with aa 422-461 of mouse ornithine decarboxylase)).

3.2   synTF construct

The synTF used for transfection in qRT-PCR experiments contains the following part of the SP1 activation domain (Residues 263 – 499) [PMID: 8278363] as previously described [9]:   NITLLPVNSVSAATLTPSSQAVTISSSGSQESGSQPVTSGTTISSASLVSSQASSSSFFTNANSY STTTTTSNMGIMNFTTSGSSGTNSQGQTPQRVSGLQGSDALNIQQNQTSGGSLQAGQQKE   GEQN-QQTQQQQILIQPQLVQGGQALQALQAAPLSGQTFTTQAISQETLQNLQLQAVPNSGP   IIIRTPTVGP-NGQVSWQTLQLQNLQVQNPQAQTITLAPMQGVSLGQTSSSN. The SP1 activation domain is fused to an N-terminal zinc-finger binding domain with a GGGGS flexible linker and expressed under control of a ubiquitin promoter containing a 5' sv40 nuclear localization sequence, C-terminal HA and rabbit globin polyA 3' UTR.

3.3   PCR plasmids

Used primer sequences are (5'-3'):
b-Actin_fwd: GGCACCCAGCACAATGAAGATCAA
b-Actin_rev: TAGAAGCATTTGCGGTGGACGATG
GAPDH_fwd: ACATCGCTCAGACACCATG
GAPDH_rev: TGTAGTTGAGGTCAATGAAGGG
eGFP_fwd: AAGTTCATCTGCACCACCG
eGFP_rev: TCCCTTGAAGAAGATGGTGCG
synTFzf_fwd: TTTTCGAGAAGACA
synTFzf_rev: GCTGCTGTGGTCGG

3.4   TFs whose motifs were checked for their presence in the designed sequences

Cebpa, Cic, Creb1, Ddit3, Elk1, Fli1, Fos, Gata1, Gata2, Gfi1, Gfi1b, Ikzf1, Irf7, Klf1, Klf4 Lyl1, Mecom, Meis1, Meis2, Myb Myc, Nfix, Nfkb1, Nfyc, Nr2c2, Pbx1, Runx1, Rxrg, Sp1, Spdef, Spi1, Stat5a, Tcf12, Tcf3, Tfap2a, Trp53, Yy1, Zbtb7a

4   Supplementary Figures

A

Linear framework graph

B

Equilibrium graph

Cycle condition

$$a_T x a_{p,\{T\}} b_{T,\{p\}} b_p = a_p a_{T,\{p\}} x b_{p,\{T\}} b_T$$

$$K_T \equiv \frac{a_T}{b_T}$$

$$K_p \equiv \frac{a_p}{b_p}$$

$$\omega \equiv \frac{a_{p,\{T\}}}{b_{p,\{T\}}} \frac{b_p}{a_p} = \frac{a_{T,\{p\}}}{b_{T,\{p\}}} \frac{b_T}{a_T}$$

$r(x) = P_3(x) + P_4(x)$

transcription rate

$r(x) = P_3(x) + P_4(x)$

transcription rate

$$r^*(x) = \frac{\begin{aligned}&b_T a_p b_{p,T} + b_T a_p b_{T,p} + a_p a_{p,T} b_{T,p} + a_T x a_{p,T} b_{T,p} + \\ & a_T x b_p a_{p,T} + a_T x a_{p,T} a_{T,p} x + a_p a_{p,T} a_{T,p} x + b_T a_p a_{T,p} x \end{aligned}}{\begin{aligned}& b_T b_{p,\{T\}} a_{T,\{p\}} x + b_T b_p b_{p,\{T\}} + b_T b_p b_{T,\{p\}} + b_p a_{p,\{T\}} b_{x,\{p\}} + \\ & a_T x b_p b_{T,p} + a_T x b_p b_{p,T} + a_T x b_{p,T} a_{T,p} x + a_p b_{p,T} a_{T,p} x + \\ & b_T a_p b_{p,T} + b_T a_p b_{T,p} + a_p a_{p,T} b_{T,p} + a_T x a_{p,T} b_{T,p} + \\ & a_T x b_p a_{p,T} + a_T x a_{p,T} a_{T,p} x + a_p a_{p,T} a_{T,p} x + b_T a_p a_{T,p} x \end{aligned}}$$

$$r^*(x) = \frac{K_p + K_p \omega K_T x}{1 + K_p + K_T(1 + \omega K_p)x}$$
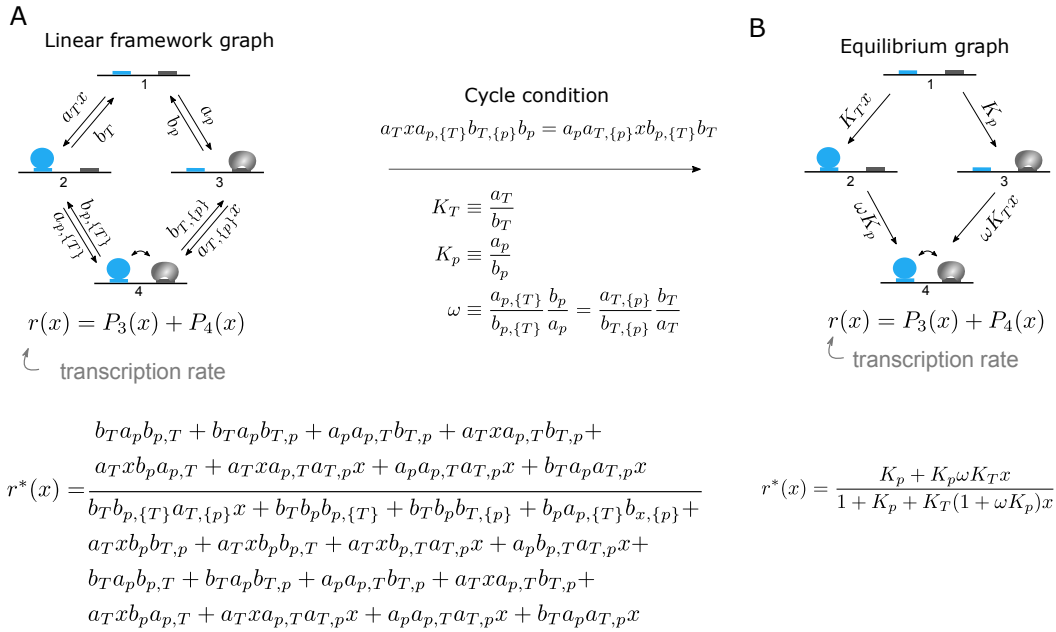
Figure S1: Example of the linear framework formalism for a simple recruitment model of polymerase. A-B) TF: blue disc. Polymerase: gray kidney bean shape. Transcription rate is the sum of probabilities of the states with polymerase bound ($q_1 = 0$, $q_2 = 0$, $q_3 = 1$, $q_4 = 1$). A) Non-equilibrium model. The steady-state transcription rate is given in terms of the graph edge labels, according to the Matrix-Tree theorem (SI Appendix). B) Equilibrium model and corresponding steady-state transcription rate.
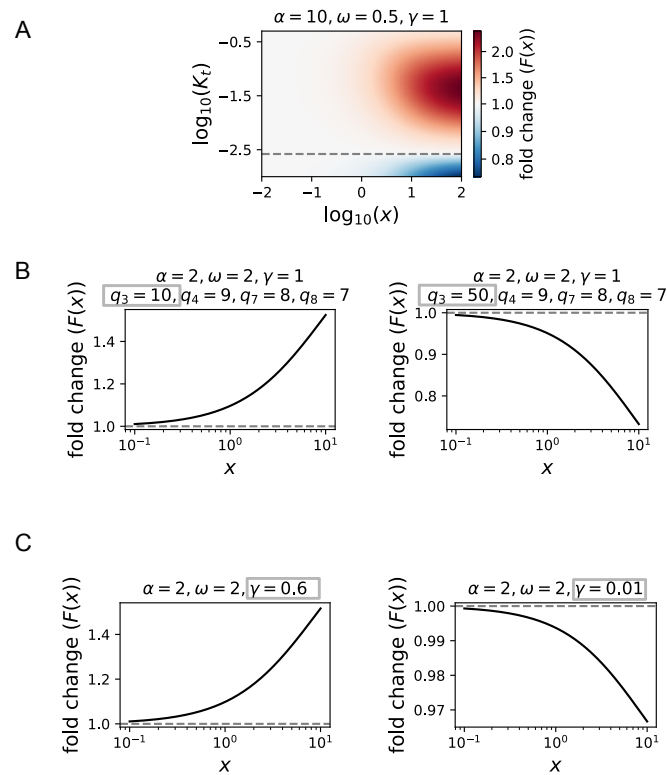
Figure S2: Duality for the equilibrium TF-chromatin model of Fig. 1C in "incoherent mode". A) Activation or repression depending on K_t. $K_T = 0.1, K_p = 0.01, \beta = 50, q_3 = 1, q_4 = 1, q_7 = 1, q_8 = 1$. B) Activation or repression depending on q_i. $K_T = 0.1, K_p = 0.01, \beta = 50, K_t = 0.005$. C) Activation or repression for inconsistent cooperativities, depending on γ. $K_T = 0.1, K_p = 0.01, \beta = 10, K_t = 0.1, q_3 = 1, q_4 = 1, q_7 = 1, q_8 = 1$.
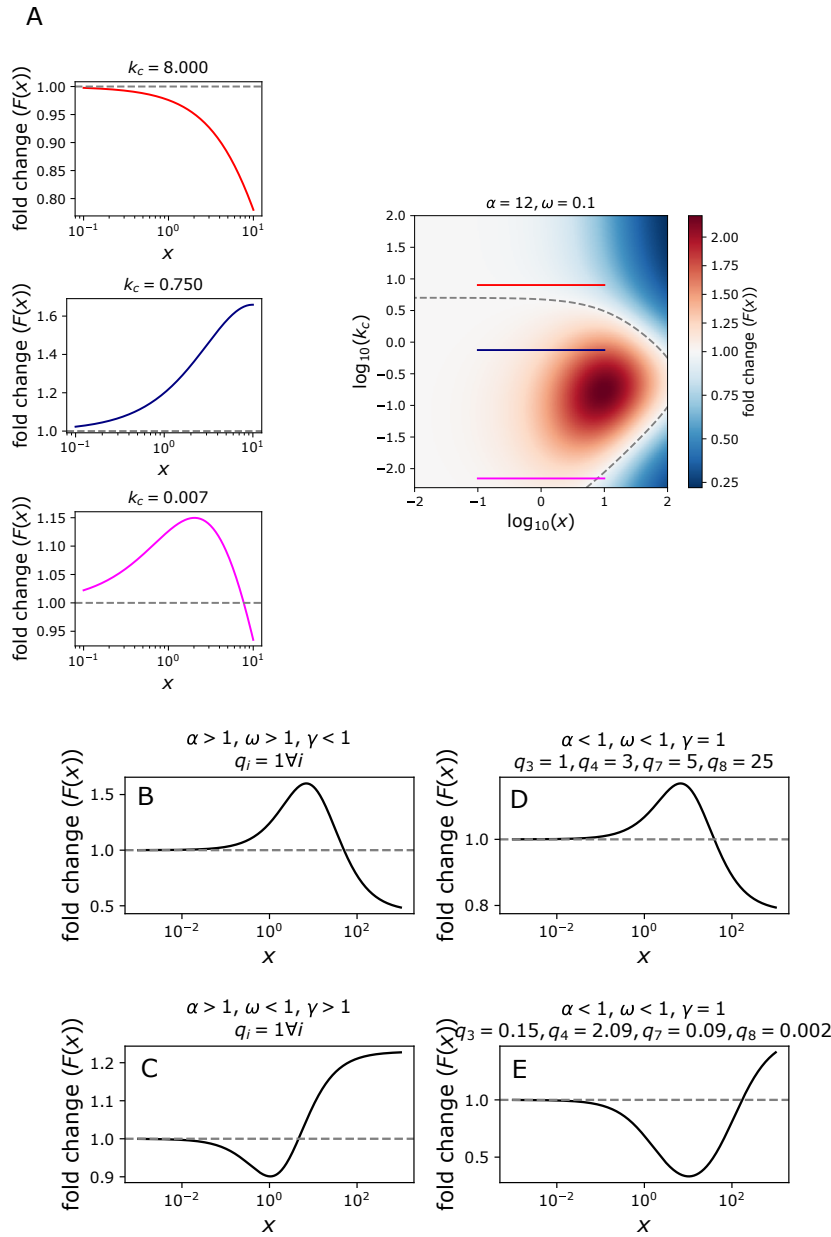
Figure S3: Non-monotonic responses for the non-equilibrium TF-chromatin model in Fig. 1E, when the TF has inconsistent effects on transcription output. A) When $\alpha > 1$ and $\omega < 1$, the basal closing rate k_c can tune the response type between activation, repression or non-monotonicity. $a_T = 1$, $b_T = 14$, $a_p = 0.1$, $b_p = 100$, $k_o = 0.01$, $\beta = 0.02$, $\gamma = 1$. B-E) Non-monotonic responses when either $\gamma \neq 1$, or $q_i \neq 1$. Note that the y-axis span different ranges. Parameter values:
B: $a_T = 1$, $b_T = 10$, $a_p = 0.1$, $b_p = 100$, $k_o = 0.01$, $k_c = 0.5$, $\alpha = 5$, $\beta = 0.001$, $\gamma = 0.01$, $\omega = 5$, $q_3 = 1$, $q_4 = 1$, $q_7 = 1$, $q_8 = 1$.
C: $a_T = 10.4$, $b_T = 18.4$, $a_p = 0.016$, $b_p = 774.4$, $k_o = 0.014$, $k_c = 3.36$, $\alpha = 8.34$, $\beta = 0.17$, $\gamma = 85.8$, $\omega = 0.07$, $q_3 = 1$, $q_4 = 1$, $q_7 = 1$, $q_8 = 1$.
D: $a_T = 1$, $b_T = 10$, $a_p = 0.1$, $b_p = 100$, $k_o = 0.01$, $k_c = 0.5$, $\alpha = 0.3$, $\beta = 0.001$, $\gamma = 1$, $\omega = 0.5$, $q_3 = 1$, $q_4 = 3$, $q_7 = 5$, $q_8 = 25$.
E: $a_T = 0.065$, $b_T = 0.1$, $a_p = 0.022$, $b_p = 0.17$, $k_o = 0.18$, $k_c = 0.0033$, $\alpha = 0.0077$, $\beta = 0.067$, $\gamma = 1$, $\omega = 0.7$, $q_3 = 0.15$, $q_4 = 2.09$, $q_7 = 0.09$, $q_8 = 0.002$.
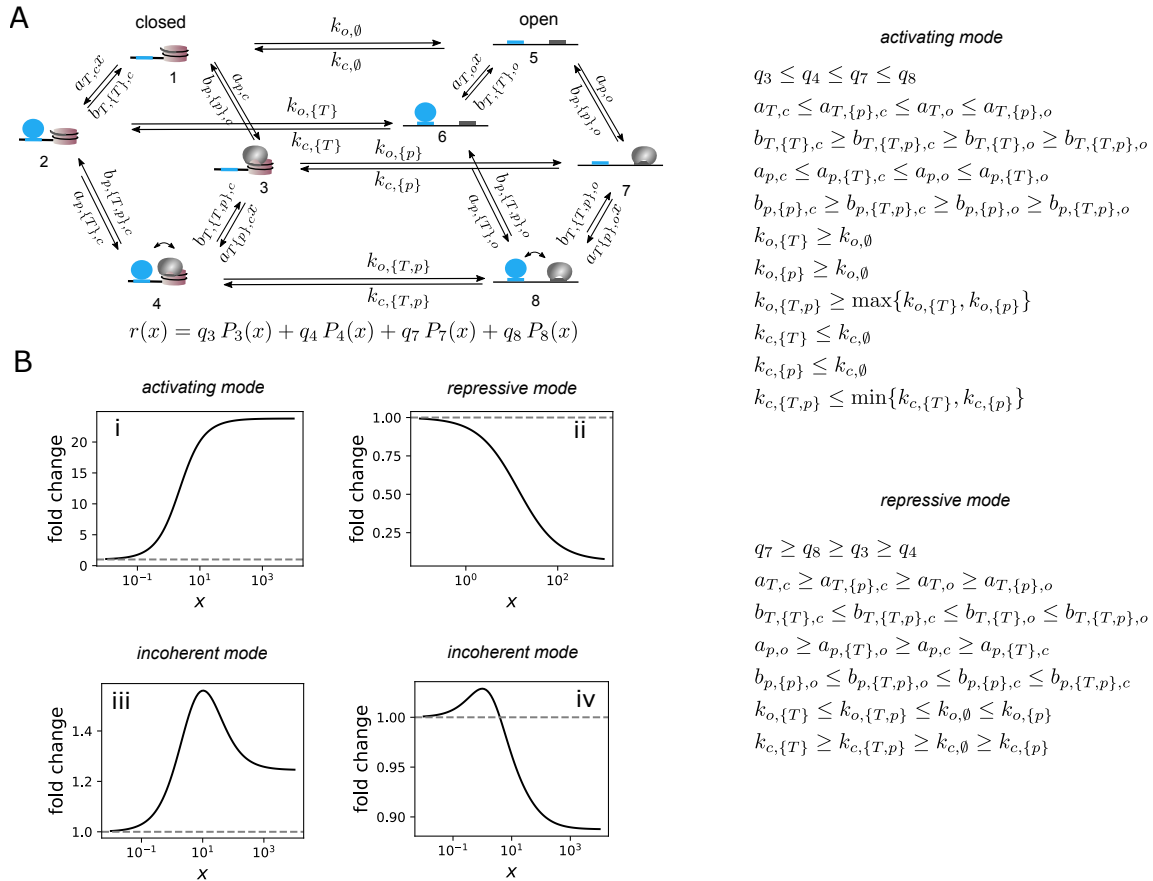
Figure S4: Non-equilibrium TF-chromatin model with arbitrary rate relationships. A) Cartoon of the model and labelling. The conditions for purely activating or repressing modes are on the right. B) Examples for parameter sets that fulfill the conditions for activation (i), repression (ii) or not (iii-iv). The parameter values are given in Table S1. Note the different axes ranges.
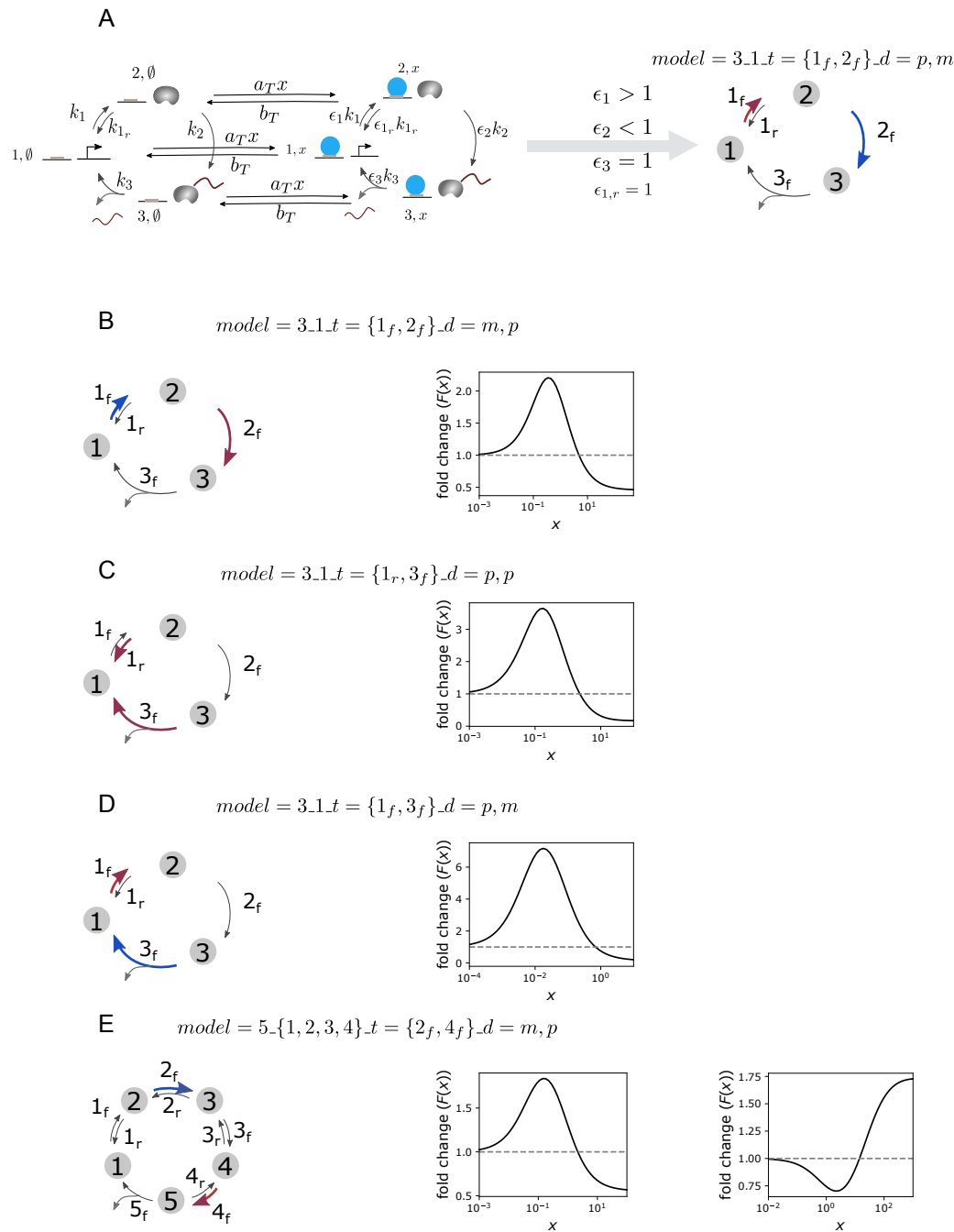
Figure S5: Generality of the non-monotonic responses in the polymerase cycle model. A) Illustration of the notation to refer to each model, as described in detail in the Supplementary text, for the model in Fig. 2C. The representation on the right is used to summarise the features of each model: number of states (3), reversible transitions (1), and transitions affected by the TF (color-coded in red for the transition accelerated by the TF, and blue for the transition slowed down. B-E) Examples of non-monotonic responses for different model implementations. Parameter values: B) $k_1 = 356.52, \epsilon_1 = 0.04, k_{1_r} = 10.67, \epsilon_{1_r} = 1.00, k_2 = 36.00, \epsilon_2 = 94.62, k_3 = 338.96, \epsilon_3 = 1.00, a_T = 658.76, b_T = 201.36$. C)$k_1 = 2855.20, \epsilon_1 = 1.00, k_{1_r} = 10000.00, \epsilon_{1_r} = 85.48, k_2 = 46.46, \epsilon_2 = 1.00, k_3 = 1.00, \epsilon_3 = 17.86, a_T = 917.20, b_T = 217.86$. D)$k_1 = 1.00, \epsilon_1 = 100.00, k_{1_r} = 14.58, \epsilon_{1_r} = 1.00, k_2 = 1.00, \epsilon_2 = 1.00, k_{2_r} = 858.59, \epsilon_{2_r} = 1.00, k_3 = 34.24, \epsilon_3 = 0.01, a_T = 10000.00, b_T = 459.72$. E)left: $k_1 = 9165.58, \epsilon_1 = 1.00, k_{1_r} = 672.27, \epsilon_{1_r} = 1.00, k_2 = 1.00, \epsilon_2 = 0.21, k_{2_r} = 4.44, \epsilon_{2_r} = 1.00, k_3 = 559.60, \epsilon_3 = 1.00, k_{3_r} = 478.39, \epsilon_{3_r} = 1.00, k_4 = 3.16, \epsilon_4 = 33.64, k_{4_r} = 7.18, \epsilon_{4_r} = 1.00, k_5 = 2304.23, \epsilon_5 = 1.00, a_T = 454.67, b_T = 280.12$. right: $k_1 = 1.00, \epsilon_1 = 1.00, k_{1_r} = 103.81, \epsilon_{1_r} = 1.00, k_2 = 10000.00, \epsilon_2 = 0.12, k_{2_r} = 781.45, \epsilon_{2_r} = 1.00, k_3 = 1.00, \epsilon_3 = 1.00, k_{3_r} = 279.75, \epsilon_{3_r} = 1.00, k_4 = 10000.00, \epsilon_4 = 100.00, k_{4_r} = 482.75, \epsilon_{4_r} = 1.00, k_5 = 1.00, \epsilon_5 = 1.00, a_T = 16.31, b_T = 26.10$.
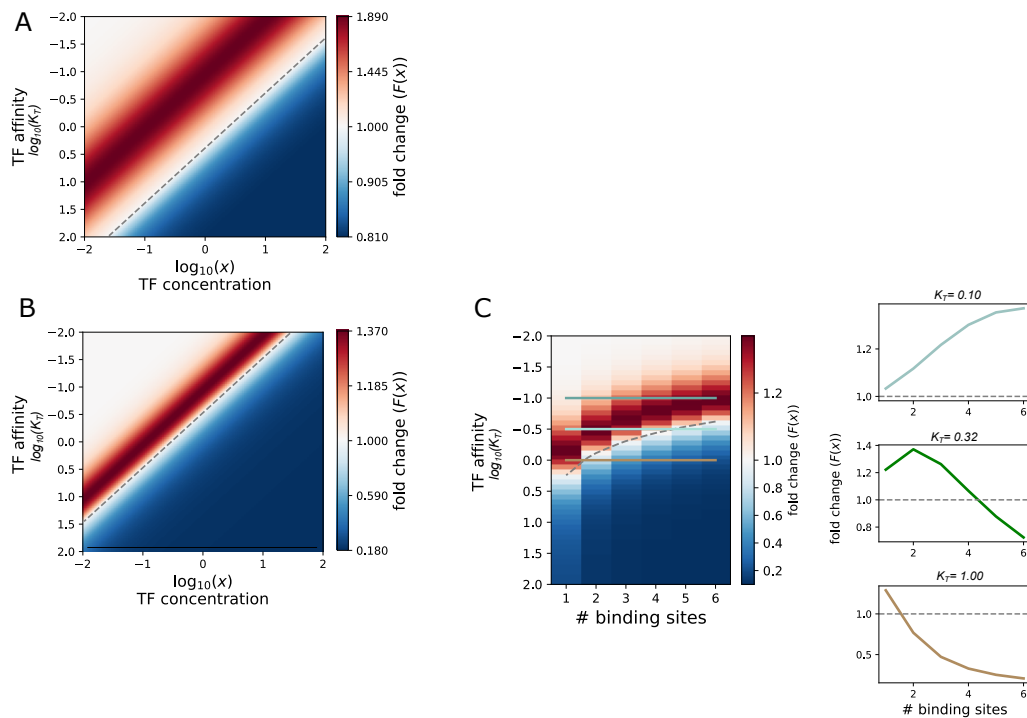
Figure S6: Regulation of the polymerase cycle model by average TF occupancy on DNA. A) Colormap showing fold change as a function of TF concentration and affinity for a TF that binds to 1 site, for the same parameter set as in Fig. 2D. B-C) Effect of increased non-linearity. Same parameter set as in A, with $h = 2$ instead of 1. B) Fold change as a function of TF concentration and affinity. C) Fold change as a function of affinty and number of binding sites, for $x = 0.173$. The lineplots on the right correspond to the response at the three affinity levels marked on the colormap.
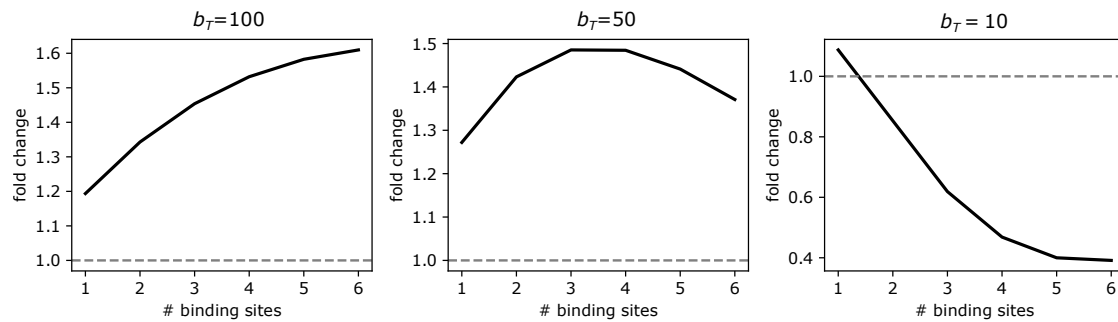
Figure S7: Affinity-dependent switch for the model in Fig. 2C extended to have more than one site. The TF, assumed at concentration 1 c.u., is assumed to bind to each site always with the same on rate $(10(c.u. \times t.u.)^{-1})$, and unbind from each site always with the same off rate as indicated on the title of each plot. When the TF is bound at k sites, it is assumed to affect the polymerase cycle rate i by factor $\epsilon_i = k f_i$. Parameter values (a.u.): $k_{1,\emptyset} = 0.5$, $k_{r1,\emptyset} = 10$, $k_{2,\emptyset} = 0.5$, $k_{3,\emptyset} = 10$, $f_1 = 5$, $f_{r1} = 1/k$ (no effect), $f_2 = 0.01$, $f_3 = 1/k$ (no effect).
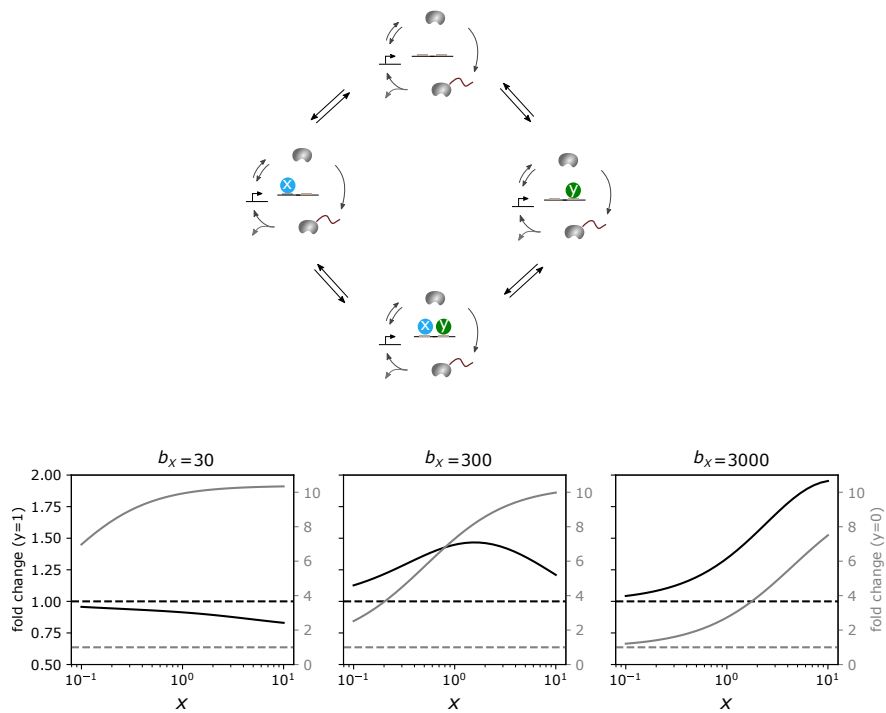
Figure S8: Expansion of the model in Fig. 2C to account for two TFs, denoted X and Y, and duality from functional interference between them. The system has a binding and unbinding reaction connecting any given cycle state with a given binding configuration, to the same state with another binding configuration, depending on whether a molecule has bound or unbound. The cartoon does not show all binding and unbinding transitions, for clarity. We assume that the binding and unbinding rate of each TF is the same regardless of which cycle state it binds to. Parameter values (same notation as in Fig. 2, with binding (unbinding) rate for TF Z denoted as a_Z (b_Z), and $k_{i,S}$ denotes the cycle transition rate i when the TFs in S are bound (a.u.): $a_X = 500$, $a_Y = 2000$, $b_Y = 300$, $k_{1,\emptyset} = 100$, $k_{1r,\emptyset} = 50$, $k_{2,\emptyset} = 25$, $k_{3,\emptyset} = 500$, $k_{1,X} = 500$, $k_{1r,X} = 50$, $k_{2,X} = 375$, $k_{3,X} = 500$, $k_{1,Y} = 100$, $k_{1r,Y} = 50$, $k_{2,Y} = 2500$, $k_{3,Y} = 500$, $k_{1,\{X,Y\}} = 500$, $k_{1r,\{X,Y\}} = 50$, $k_{2,\{X,Y\}} = 50$, $k_{3,\{X,Y\}} = 500$.
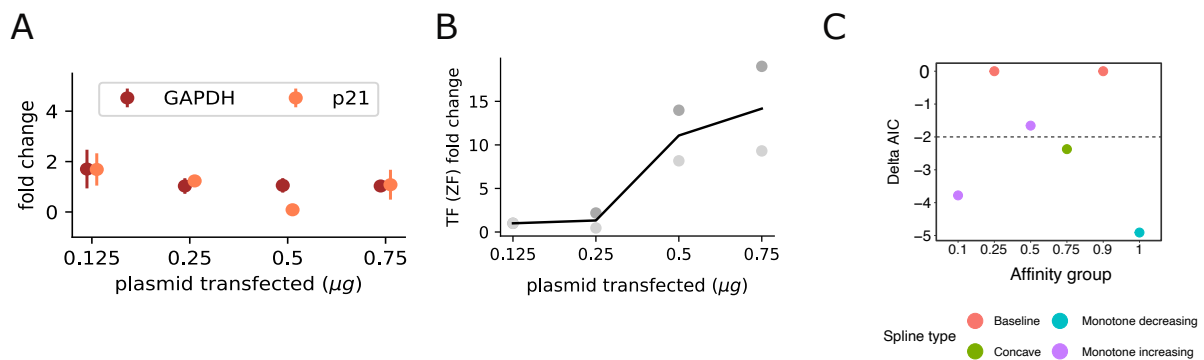
Figure S9: Supplementary experimental panels. A) mRNA fold change for GAPDH and p21 over the range of synTF transfected concentrations used in these experiments, for which expression of these genes remains approximately constant, as measured by qPCR. B) TF (ZF) mRNA fold change, relative to the lowest input condition in the plot, as measured by qPCR. Data corresponding to two biological repeats is shown as a scatter plot with two shades of gray, with the black line representing the mean. C) Lowest AIC value difference of the fits of the data in fig. 4E with different shape constraints on P-splines, relative to a baseline horizontal fit. The dotted line set at -2 indicates the consensus for when a model is considered to be (moderately) significantly better than another [10]. This analysis supports an affinity-dependent switch in the response trend, with monotonically-increasing response at the lowest affinity (0.1), concave (non-monotonic) at 0.75, and monotone decreasing at the highest affinity.

Table S1: Parameters for the plots in Fig. S4.

parameter	i	ii	iii	iv
$a_{T,c}$	0.50	0.08	0.50	0.50
$a_{T,o}$	0.94	0.02	0.94	0.94
$b_{T,\{T\},c}$	2.00	1.00	2.00	2.00
$b_{T,\{T\},o}$	0.75	4.00	0.75	0.75
$a_{p,c}$	0.01	0.01	0.01	0.01
$a_{p,o}$	0.02	0.04	0.01	0.02
$b_{p,\{p\},c}$	100.00	60.00	10.00	100.00
$b_{p,\{p\},o}$	37.50	10.00	2.50	37.50
$a_{p,\{T\},c}$	0.01	0.01	0.00	0.01
$a_{p,\{T\},o}$	0.04	0.02	0.00	0.04
$b_{p,\{T,p\},c}$	75.00	180.00	10.00	75.00
$b_{p,\{T,p\},o}$	18.75	20.00	12.50	18.75
$a_{T,\{p\},c}$	0.62	0.04	0.62	0.62
$a_{T,\{p\},o}$	0.94	0.01	0.94	0.94
$b_{T,\{T,p\},c}$	1.50	2.00	1.50	1.50
$b_{T,\{T,p\},o}$	0.38	8.00	0.38	0.38
$k_{o,\emptyset}$	0.10	0.10	0.10	0.50
$k_{c,\emptyset}$	1.00	1.00	50.00	1.00
$k_{o,\{T\}}$	1.00	0.03	10.00	5.00
$k_{c,\{T\}}$	0.10	4.00	25.00	1.00
$k_{o,\{p\}}$	0.10	0.20	0.10	0.50
$k_{c,\{p\}}$	1.00	0.50	50.00	1.00
$k_{o,\{T,p\}}$	1.00	0.05	10.00	5.00
$k_{c,\{T,p\}}$	0.10	2.00	25.00	1.00
q_3	1.00	1.00	0.10	0.25
q_4	1.00	1.00	0.50	0.00
q_7	1.50	1.00	0.50	1.00
q_8	2.00	1.00	1.00	0.10

Supplementary References

[1] Jeremy Gunawardena. "A Linear Framework for Time-Scale Separation in Nonlinear Biochemical Systems". In: *PLoS ONE* 7.5 (2012), e36321.

[2] Inomzhon Mirzaev and Jeremy Gunawardena. "Laplacian Dynamics on General Graphs". In: *Bulletin of Mathematical Biology* 75.11 (2013), pp. 2118–2149.

[3] Tobias Ahsendorf et al. "A framework for modelling gene regulation which accommodates non-equilibrium mechanisms". In: *BMC Biology* 12.1 (2014), p. 102.

[4] Javier Estrada et al. "Information Integration and Energy Expenditure in Gene Regulation". In: *Cell* 166.1 (2016), pp. 234–244.

[5] Kee-Myoung Nam, Rosa Martinez-Corral, and Jeremy Gunawardena. "The linear framework: using graph theory to reveal the algebra and thermodynamics of biomolecular systems". en. In: *Interface Focus* 12.4 (2022), p. 20220013.

[6] Gary K Ackers, A. D. Johnson, and Madeline A Shea. "Quantitative model for gene regulation by lambda phage repressor." In: *Proceedings of the National Academy of Sciences* 79.4 (1982), pp. 1129–1133.

[7] Lacramioara Bintu et al. "Transcriptional regulation by the numbers: models." In: *Current opinion in genetics & development* 15.2 (2005), pp. 116–24.

[8] Rosa Martinez-Corral et al. "The Hill function is the universal Hopfield barrier for sharpness of input–output responses". In: *Proceedings of the National Academy of Sciences* 121.22 (2024), e2318329121.

[9] Rosa Martinez-Corral et al. "Transcriptional kinetic synergy: A complex landscape revealed by integrating modeling and synthetic biology". en. In: *Cell Syst* 14.4 (2023), 324–339.e7.

[10] M.C. Greenwood. *Intermediate Statistics with R*. Open textbook library. Montana State University, 2021.