

Integrating regulatory information via combinatorial control of the transcription cycle

Clarissa Scholes^{1,2}, Angela H. DePace^{1,*} and Álvaro Sánchez^{3,*}

¹ Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

² Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA.

³ The Rowland Institute at Harvard, Harvard University, Cambridge, MA 02142, USA.

* These authors contributed equally to this work

Correspondence: A.H.D., angela_depace@hms.harvard.edu; A.S., sanchez@rowland.harvard.edu

ABSTRACT

Combinatorial regulation of gene expression by multiple transcription factors (TFs) enables cells to carry out sophisticated computations that are key to cellular decision-making. How is the information contained in multiple TF binding sites integrated to dictate the rate of transcription? The dominant model is that direct or indirect physical interactions between TFs enable them to combinatorially recruit each other and RNA polymerase to the promoter. Here we develop a quantitative framework to explore an alternative model, where combinatorial gene regulation can result from TFs working on different kinetic steps of the transcription cycle. Our results clarify the null hypotheses for independent action of TFs and show that combinatorial control of the transcription cycle can generate a wide range of analog and Boolean computations without requiring the input regulators to be simultaneously co-localized in the nucleus. This work emphasizes the importance of deciphering TF function beyond activation and repression, highlights the role of the basal promoter in processing regulatory information and suggests qualitative explanations for the flexibility of regulatory evolution.

keywords: transcription; kinetics; combinatorial control; frequency modulation; quantitative modeling; boolean computations

INTRODUCTION

To control gene expression, multiple transcription factors (TFs) bind to clusters of their cognate binding sites in regulatory DNA; this relays information from environmental and developmental cues into gene expression outcomes. Combinatorial regulation enables a small number of TFs to implement different types of computations (Struhl 1991; Carey 1998) in different cell types or over developmental time (Spitz & Furlong 2012). A grand challenge is to decipher how information contained in multiple binding sites—or more broadly, in multiple regulatory elements in a gene locus—is integrated to quantitatively determine the rate of transcription (Figure 1A,B).

To analyze how computations are carried out by TFs binding to regulatory DNA, computational models of combinatorial gene regulation have largely focused on how TFs interact, cooperatively or competitively, to

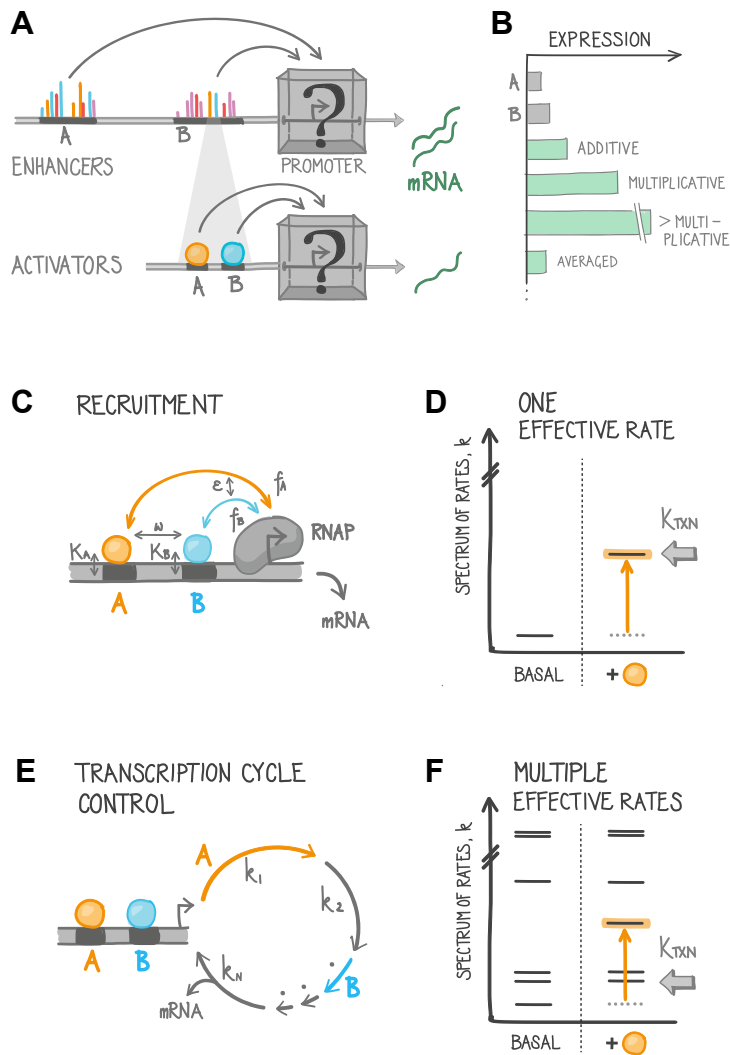


Figure 1. Promoters integrate information from multiple regulatory sequences

(A) Multiple regulatory sequences (top) or individual bound transcriptional activators (bottom) impinge on a given promoter at the same time. Together these regulatory sequences determine the level and timing of gene expression. We ask how these inputs are integrated by the promoter, by defining the overall expression level driven by two elements (either enhancers or activators) as a function of their individual effects.

(B) Two regulatory elements, A and B, may combine in different ways to determine the overall expression level.

(C) Models of combinatorial transcriptional control usually assume that the single limiting step is "recruitment" of RNA Polymerase (RNAP) to the promoter. Recruitment is mediated by physical interactions between different transcription factors (TFs, A and B) and the DNA, and between TFs and RNAP. The rate of transcription is assumed to be proportional to RNAP occupancy of the promoter. K_A , dissociation constant for TF A. f_A , strength of interaction between TF A and RNAP. ω , cooperativity between TFs A and B in binding to DNA. ϵ , cooperativity between the TFs in recruiting RNAP.

(D) Recruitment models implicitly assume a single slow effective rate in transcription. Transcription is limited by the dissociation of RNAP from the promoter, and activators work by reducing this dissociation rate. The single effective rate is described by the rate constant k (first column). Adding an activator (orange) accelerates this effective rate, raising the overall rate of transcription (K_{TXN} arrow). All other steps in transcription are assumed to be infinitely fast (not shown).

(E) The transcription cycle contains numerous distinct steps (Box 1). Shown here are the effective rates of N arbitrary steps, with shorter arrows indicating faster rates. TFs may work on the same or different steps in the cycle. Here, TF A accelerates effective rate k_1 (orange) while B accelerates one of the faster, unlabeled rates (blue).

(F) The effective rates of the transcription cycle are depicted before and after addition of an activator (orange) that works on the slowest step. The overall rate of transcription increases when the activator is added, but is limited by the next slowest rate.

recruit or repel one another and the basal transcriptional machinery (Bintu et al. 2005a; He et al. 2010; Veitia 2003; Coulon et al. 2013; Wang et al. 1999; Frank et al. 2012; Gertz et al. 2009). A key implicit assumption of these “recruitment” models (Ptashne & Gann 1997) is that there is a single effective rate-limiting step in transcription (Figure 1D). It follows from this assumption that all TFs act by modulating this rate-limiting step. At many bacterial genes, where RNA polymerase (RNAP) is in direct competition with a repressor whose binding site overlaps its own, the approximation of a single rate limiting step may be reasonable (Garcia & Phillips 2011; Brewster et al. 2012). It should be noted, however, that this is just an approximation, and that many bacterial promoters do not conform to it (Box 1). A frequent additional assumption is that TF and RNAP binding can be calculated at equilibrium, thus this class of models is often referred to as “thermodynamic” (Bintu et al. 2005a). Several recent reviews give practical guidelines on building and applying these models (Bintu et al. 2005a; Bintu et al. 2005b; Sherman & Cohen 2012; Dresch et al. 2013).

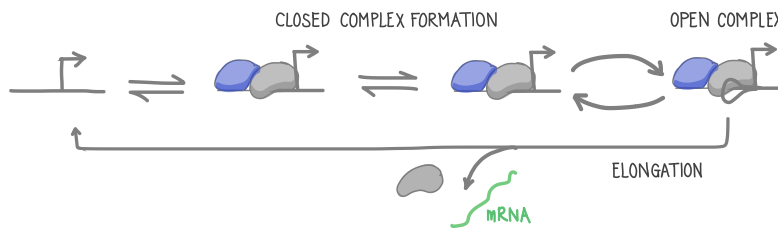
The assumption that there is only a single rate-limiting step in transcription may not, however, be appropriate for eukaryotes. The eukaryotic transcription cycle is composed of numerous distinct steps, from clearing of nucleosomes around the promoter to release of paused RNAP (Box 1), and various of these steps can be rate-limiting (Fuda et al. 2009; Green 2005; Nechaev & Adelman 2011; Jonkers et al. 2014; Hahn 2004; Choubey et al. 2015). Analysis of transcriptional dynamics in eukaryotic cells indicate that there are multiple slow biochemical reactions required to turn a gene on, and that at the promoters studied their rates must be within the same order of magnitude (Coulon et al. 2013; Zoller et al. 2015; Choubey et al. 2015).

If two or more steps have similar rates, control of transcription might not be dominated by regulation of a single step (Figure 1F); this could allow for combinatorial control by regulators working on different steps, as proposed over 20 years ago by Herschlag & Johnson (1993). Indeed, TFs are known to work on distinct steps in the transcription cycle (Fuda et al. 2009; Blau et al. 1996; Govind et al. 2005; Stasevich et al. 2014; Yankulov et al. 1994; Brown et al. 1998; Liu et al. 2003). For example, the activator SP1 works primarily on initiation (Blau et al. 1996) and Tat targets elongation (Barboric & Peterlin 2005); many (perhaps most) TFs affect multiple steps (e.g. Gcn4 (Govind et al. 2005) and CRP (Liu et al. 2003)). Since it was proposed, some evidence consistent with kinetic control of the transcription cycle has come from a small number of experimental studies (Blau et al. 1996; Gonzalez-Couto et al. 1997; Keung et al. 2014; Blair et al. 1996). For example, Bentley and colleagues saw greater-than-additive expression only between pairs of activation domains that worked on different steps (Blau et al. 1996).

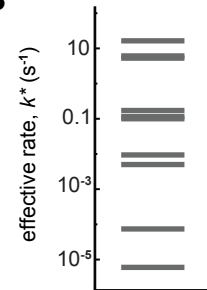
Here, we aim to provide a synthesis of these theoretical and experimental observations by exploring combinatorial control via regulation of different slow steps in the transcription cycle. We emphasize that this type of kinetic control is compatible with combinatorial control arising from cooperative interactions between TFs in recruiting RNAP; they are not mutually exclusive mechanisms. For the sake of clarity, here we focus on a quantitative model of kinetic control as this mechanism is relatively under explored. Quantitative models enable us to reason about complex processes and to make falsifiable predictions that can be tested experimentally (Gunawardena 2014; Phillips 2015). For transcription, such models are also essential for drawing mechanistic interpretations from increasing volumes of high-throughput quantitative data that attempt to empirically characterize regulatory DNA (e.g. Sharon et al. 2012; Smith et al. 2013; Patwardhan et al. 2012). Our goal is here is to probe the regulatory outcomes and computations that

emerge when cells combinatorially regulate different kinetic steps. We show that kinetic control of the transcription cycle can be used to implement a wide range of logical and analog computations, without requiring any cooperative interactions between input TFs. The pairwise interactions between TFs that are needed to produce many of these computations under recruitment-based models require that the TFs are simultaneously co-localized in the nucleus; combinatorial control of the transcription cycle can recapitulate these computations without the requirement for TF co-localization.

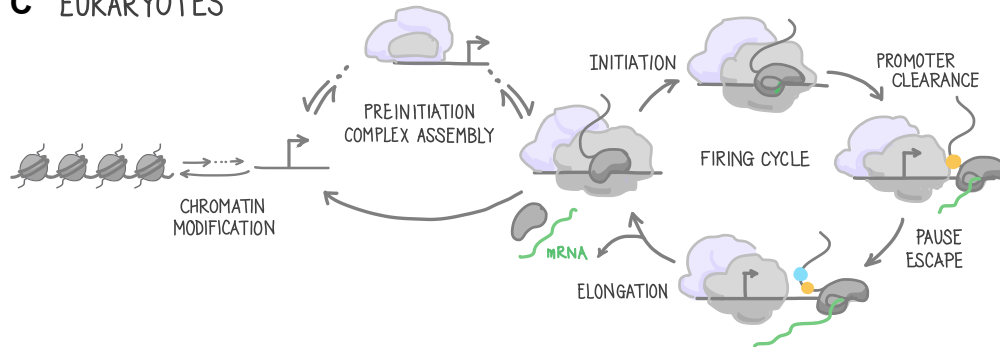
A PROKARYOTES



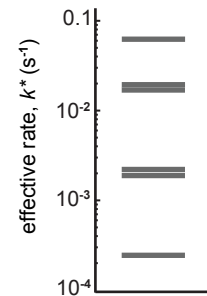
B



C EUKARYOTES



D



Box 1. The transcription cycle in prokaryotes and eukaryotes

Prokaryotes

(A) RNAP is recruited to the promoter through association with a sigma factor to form the closed complex (Gross et al. 1998; Browning & Busby 2004). Subsequent melting of the DNA forms the open complex (McClure 1985) and RNAP then escapes the promoter into elongation. It was widely assumed that recruitment was the rate-limiting step in bacterial transcription (Ptashne & Gann 1997); however, the transition from initiation to elongation is rate limiting at many promoters (Reppas et al. 2006; Wade & Struhl 2008). Diagram of a bacterial transcription cycle adapted from Friedman & Gelles (2012).

(B) We used the forward and reverse reaction rates measured in Friedman & Gelles (2012) to calculate the “effective” forward rates through the transcription cycle (Results and Figure S1A-C). As this rates diagram shows, assuming a single slow step in transcription may be problematic even for bacterial promoters: the two slowest effective rates are within an order of magnitude of one another.

Eukaryotes

(C) The transcription cycle begins with the formation of the preinitiation complex comprised of RNAP and general TFs (Fuda et al. 2009). The chromatin surrounding the promoter is modified, which is coordinated by DNA-bound TFs that recruit chromatin remodelers and nucleosome modifying enzymes. Transcription initiation can take place once the DNA has been unwound to form the open complex. RNAP must subsequently escape the grip of proteins at the promoter to enter into elongation. At most genes in higher eukaryotes, RNAP pauses downstream of the promoter and must be actively released from this paused state into productive elongation through the body of the gene (reviewed in Zeitlinger, Stark, et al. 2007; Jonkers et al. 2014; Adelman & Lis 2012). Diagram of higher eukaryotic transcription cycle adapted from Fuda et al. (2009) and Boettiger 2013.

(D) The “effective” rates (see Results) for a simplified eukaryotic transcription cycle were calculated from Stasevich et al (2014) (Figure S1D-F).

RESULTS

Calculating transcription from a cycle with no internal loops

To give intuition for how regulation of a transcription cycle is affected by all of its component steps, we first derive a closed form expression for a cycle with no internal loops. Here, the rate of transcription is the inverse of the sum of the transition times between states, but when reversible steps are included, the *effective time* taken to complete each step depends on the reverse as well as the forward transition times. We summarise the effect of a reverse rate between *state j* and *state j-1* by introducing a parameter that captures the average number of times the promoter switches backwards; we call this the number of ‘Rejections’ from *state j*,

$$R_j = \frac{k_{j,j-1}}{k_{j,j+1}}, \quad (\text{Equation 1})$$

where $k_{j,j+1}$ is the rate of the forward reaction from *state j* to the subsequent state, $j+1$, and $k_{j,j-1}$ is the rate of the backward reaction from *state j* to the previous state, $j-1$. The effective time taken to complete step j (i.e. the “handling time” of that state) depends on the number of Rejections from all of the *downstream* reversible transitions in the transcription cycle. For example, for a cycle comprised of $N = 4$ states in which the first three are connected by reversible transitions and the final step in which mRNA is produced is irreversible, we calculate the effective transition times, denoted by T_j^* for $j = 2:N$ as follows, where τ_j is the time taken to transition between states j and $j+1$:

$$\begin{aligned} T_1^* &= \tau_1 (1 + R_2 + R_2R_3 + R_2R_3R_4), \\ T_2^* &= \tau_2 (1 + R_3 + R_3R_4), \\ T_3^* &= \tau_3 (1 + R_4), \\ T_4^* &= \tau_4 (1), \\ T_{TOTAL} &= \sum_{i=1}^N T_i^*. \end{aligned} \quad (\text{Equation 2})$$

The total time taken to complete a cycle with reversible and irreversible steps (Equation 2) can be written in compact form for a cycle with an arbitrary number of steps, N , as:

$$T_{TOTAL} = \sum_{m=1}^N \tau_m \left(1 + \sum_{s=m+1}^N \prod_{j=m+1}^s R_j \right). \quad (\text{Equation 3})$$

The rate of transcription is the inverse of T_{TOTAL} . For the previous example of a four-step cycle, what we end up with is actually ten terms (all the terms in the Equation 3) that must be summed to calculate the total transcription time. This means that there are ten “effective” or “pseudo”-rates. We label these as such because they do not necessarily correspond to actual biochemical reactions but rather could be composed of both the forward and reverse rates, or multiple sub-steps. For example, the effective time the promoter spends trying to move past the first state in the cycle, T_1^* is the sum of four ‘pseudo’ times τ_1 , τ_1R_2 , $\tau_1R_2R_3$ and $\tau_1R_2R_3R_4$, which is equivalent to saying that there are four irreversible reactions in series that together comprise the transition from state 1 to state 2 (see example cycle in Figure S1A-C and Supplementary Information).

These equations give an intuition for how the output of a cycle depends its structure. Most importantly, the overall rate of transcription is not simply equivalent to the rate limiting step. Instead, the total transcription time depends on the number and reversibility of all steps in the cycle, and the time spent in any given state depends on the subsequent effective rates. An analogy would be trying to file paperwork in the context of a bureaucratic hierarchy; the number of times you will have the pleasure of meeting civil servant 1 depends on the number of times civil servants 2, 3 and 4 tell you that you've filled out your paperwork incorrectly and must try again. This is reflected in the concept of "effective" rates, which account for the structure of the transcription cycle; we thus use effective rates in the formulation of our model below.

A model based on stochastic chemical kinetics captures both reversible and irreversible steps in transcription

We develop a quantitative model of combinatorial regulation of the transcription cycle based on chemical kinetics. This formalism that can accommodate cycles of arbitrary complexity in terms of connectivity and number of steps; the overall rate of transcription is calculated using a master equation formalism that generates closed-form analytical expressions for the mean and variance of the mRNA distributions (Sanchez & Kondev 2008). Using this formalism one can model the transcription cycle as a series of $N-1$ inactive promoter states leading to state N , from which an mRNA molecule is produced as the promoter switches back to state 1, from which the cycle starts again. The promoter transitions stochastically between states N and $N+1$ with rate $k_{N,N+1}$. We explain this formal model in detail in the Experimental Procedures.

We focus here for illustration purposes on the simple case of combinatorial regulation of a transcription cycle composed of two effective forward rates (k_1 and k_2). Note that this simple case does not require the matrix formalism in order to do the appropriate calculations. However, the insights we derive from this simple case hold true even for more complicated cycles, which can be considered in the matrix formalism above.

In addition to defining the structure of the cycle, we must define how activators affect the component rates, which we define as an "enhancement" over the basal rate. In the absence of activators, the reaction proceeds at a "basal" rate, $k_{TXN}(0) = k_{basal}$. To understand the influence of activator A on this basal rate, consider a cycle with a single slow step. In the presence of activator A, this rate increases to $k_{TXN}(A) = k_{basal} + \Delta k(A)$, where $\Delta k(A)$ represents the rate enhancement (the increase in rate k) due to TF A. Similarly, $k_{TXN}(B) = k_{basal} + \Delta k(B)$. If both A and B are both present and they act independently by accelerating the same single rate in parallel (i.e. each does so through a different, non-interacting mechanism), then the rate enhancements will sum together: $\Delta k(A,B) = \Delta k(A) + \Delta k(B)$. By defining $a = \frac{\Delta k(A)}{k_{basal}}$, and $b = \frac{\Delta k(B)}{k_{basal}}$, we can express $k_{TXN}(A) = k_{basal}(1 + a)$, and similarly, $k_{TXN}(B) = k_{basal}(1 + b)$. Therefore, $k_{TXN}(A,B) = k_{basal}(1 + a + b)$, and this defines how in our model, two activators both independently work on the same step.

Now we consider how activators work on the same step, but in the context of a two step cycle. We use k_1 and k_2 to denote the basal rates of the transitions between state $1 \rightarrow 2$ and $2 \rightarrow 1$, respectively, while a_1 and b_1 denote the enhancement of rate k_1 rates by TF A and B, and likewise for rate k_2 ($a_i \geq 0$ for an activator). Thus, the rate of transcription from a two step cycle in the absence of TFs is

$$k_{TXN}(0) = \frac{k_1 \times k_2}{k_1 + k_2}, \quad (\text{Equation 4})$$

and in the presence of both TFs A and B is

$$k_{TXN}(A,B) = \frac{(1 + a_1 + b_1) k_1 \times (1 + a_2 + b_2) k_2}{(1 + a_1 + b_1) k_1 + (1 + a_2 + b_2) k_2}. \quad (\text{Equation 5})$$

This formulation applies to activators; we use a different formulation for repressors, as described in relation to Figure 2 (see also Experimental Procedures).

In the remainder of the results we consider computations that are possible when two TFs act on either or both steps. We assume throughout that TFs work entirely independently of one another; that is, that neither TF helps or hinders the activity of the other. The magnitude of enhancement that an activator has on a given step is determined both by its intrinsic strength and by its concentration. When we randomize the enhancement factor for TFs in our simulations below, this represents randomization of both of these components; thus, we do not explicitly comment on the effect of changing TF concentration alone.

Logical computations can be achieved by controlling the transcription cycle

Boolean (or logical) computations are commonly used to represent gene regulatory circuits in development and engineering because they produce sharp, on/off decisions (Greber & Fussenegger 2007; Istrail & Davidson 2005; Davidson 2011). Given the importance of these computations at the systems level, it is key to understand how these binary computations may be executed by the cell. Seminal work by Hwa and colleagues addressed this question and found that Boolean logic gates can be implemented for gene regulation by combinatorial recruitment of RNAP (Buchler et al. 2003). These authors demonstrated that all of the elemental two-input logic gates (e.g. AND, OR, NAND, NOR, EQ, XOR) can in principle be built by cooperative pairwise physical interactions between proteins; more complex gates arise by modular arrangement of these pairwise interactions. Here we explore whether logic gates can also be built by regulating the transcription cycle without requiring any physical interactions between TFs, i.e. when activators act independently of one another. Here we use “independent” to mean that the activators neither help nor hinder one another in acting on their shared promoter, in contrast other its alternate usage to describe regulators acting in distinct pathways (i.e. on different promoters).

We calculated the overall rate of transcription by two TFs individually and together, while altering the relative effective rates of the two steps and the strength and direction of TF action (activating or repressing; Figure 2, and Tables S1,2). While there are various ways in which repression could occur mechanistically, here we made the assumption that repressors act by increasing the free energy required to achieve a transition between states; as such, when more than one repressor acts on the same step, their combined effect on that rate is multiplicative. Our analysis shows that the six two-input logic gates described by Buchler et al. (2003) can also be achieved kinetically without invoking any cooperation between TFs. In the simplest example, the AND gate is generated if high expression is possible only when both TFs are present; for this to occur, TFs A and B work as activators each on a different slow step (Figure 2C). In a more complex example, the implementation of the EQ and XOR gates require each TF to activate one step while repressing the other (Figure 2G,H). This dual activation-inhibition effect might for

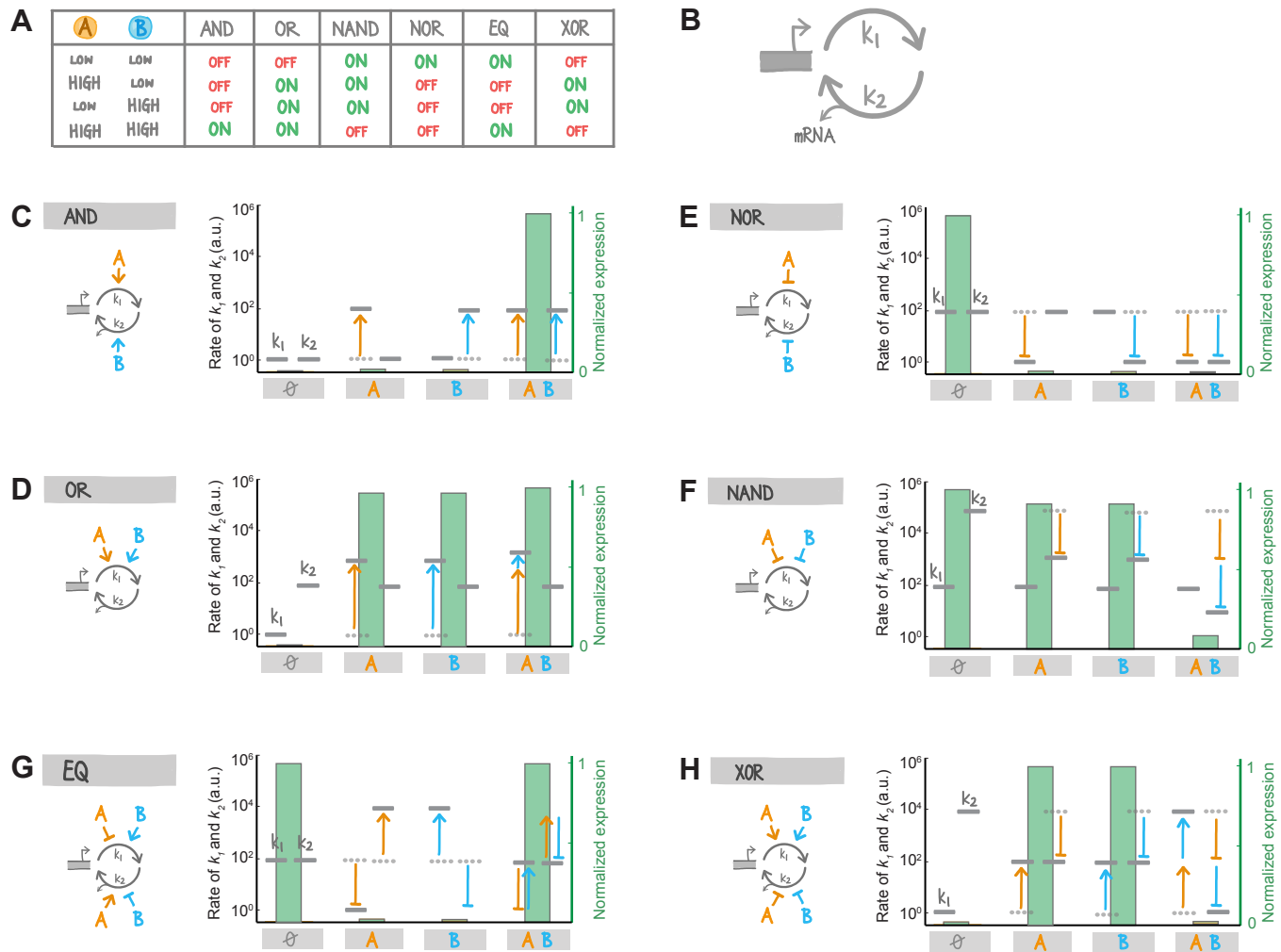


Figure 2. Logical computations could be explained by regulation of the transcription cycle

(A) Summary table showing how combinations of TFs A and B map to common Boolean or logical computations. The input TFs are either at low or high concentration, and the output is binary: the gene is ON (highly expressed) or OFF (very lowly expressed, or not expressed at all).

(B) We consider a simple two-step transcription cycle. The effective forward rates are characterized by the rate constants k_1 and k_2 , and a molecule of mRNA is produced from step two.

(C-H) Control of a simple two-step transcription cycle by two TFs can produce all of the basic Boolean computations. For each logic gate we show an example of how TFs A and B could bring about the computation by acting on one or both rates (grey horizontal bars), either as activators (arrows) or repressors (blunted arrows). The resulting expression output due to each TF acting individually or in combination is shown in green, normalised to the maximum expression in each plot such that 0 is "OFF" and 1 is "ON"; equations used to calculate expression in each case are in the Supplementary Information. For the NAND, EQ and XOR gates we assume that when TFs act to reduce the effective rate of their target step, they do so by increasing the free energy of transition; as such, their effects on the rate are multiplicative (see Supplementary Information). The effect of repressor R on a rate is described here as $k_i(R) = k_i(0)(1+r_i)$, where r_i is the ratio between the repressor concentration and the operator's dissociation constant ($r_i(0)$).

instance occur if, while accelerating one rate, a TF sterically hinders another process in the transcription cycle from taking place.

Combinatorial control of the transcription cycle allows digital computations by out-of-sync nuclear localization events

Recent experiments demonstrate that TFs can regulate genes by transient bursts of nuclear localization (Cai et al. 2008; Dalal et al. 2014; Hao & O’Shea 2012; Batchelor et al. 2011; Lahav et al. 2004). Furthermore, the timing of colocalization bursts by different regulators may be a means to combinatorially control transcription (Lin et al. 2015). Under the “recruitment” assumption, combinatorially acting TFs must be simultaneously colocalized to the nucleus (i.e. localized in-sync) in order to co-recruit one another and/or the transcriptional machinery (Figure 3A,B) (Lin et al. 2015). In contrast, control of the transcription cycle could enable combinatorial gene regulation even when the bursts of TF nuclear localization are out of sync (Figure 3C, D). All of the logical gates discussed above can be implemented by TFs that are not present in the nucleus at the same time.

The ability of combinatorial transcription cycle control to work with out-of-sync bursts of TF nuclear localization might provide a means for cells to discriminate the timing of signals. For example, consider two regulators, A and B, that combinatorially activate transcription by acting on a transcription cycle that consists of four slow steps (Figure 3C). TF A accelerates the first step of the cycle (e.g. eviction of a nucleosome) whereas B accelerates the last step (e.g. release of paused RNAP). Two intermediate steps proceed slowly but are unaffected by A or B. TFs A and B are activated by signaling pathways and transiently enter the nucleus in bursts that are separated by time τ . Although the nuclear localization bursts may have different and complex shapes (Cai et al. 2008), for simplicity and analytical tractability we model them as step functions of height k_A and k_B and duration T_A and T_B . We further assume that $k_A T_A$, $k_B T_B \gg 1$, such that their respective kinetic step can be completed during nuclear localization of each TF; we also assume the two intermediate steps have equal rates ($k_{23} = k_{34} \equiv k_i$) that are slow compared to the activated rates, such that $k_i T \ll 1$. Qualitatively, none of our conclusions depend on these assumptions, but they simplify the math, allowing us to determine the fraction of cells that respond to this time-encoded AND gate, as a function of the time between the two signals (τ) as:

$$f_{\text{respond}}(\tau) \approx 1 - e^{-k_i(T+\tau)} - k_i(T+\tau)e^{-k_i(T+\tau)}. \quad (\text{Equation 6})$$

Equation 6 is plotted in Figure 3C. For two temporally-overlapping signals, the cell will not respond to the nuclear localization events Figure (3C,D). Only when the nuclear localization bursts of A and B are shifted by a time larger than $\tau_{\text{crit}} = 1/k_i$ is a significant transcriptional response observed, producing an AND gate. Similarly, if the order of the signals is inverted, and B translocates to the nucleus before A does, the time integrated AND gate produces no transcriptional response (Figure 3D).

This simple example illustrates how kinetic control of the transcription cycle provides a straightforward mechanism by which cells could to add a time component to their logical computations. Combinatorial transcription cycle control could also enable cells to respond to the relative timing between signals (Lin et al. 2015).

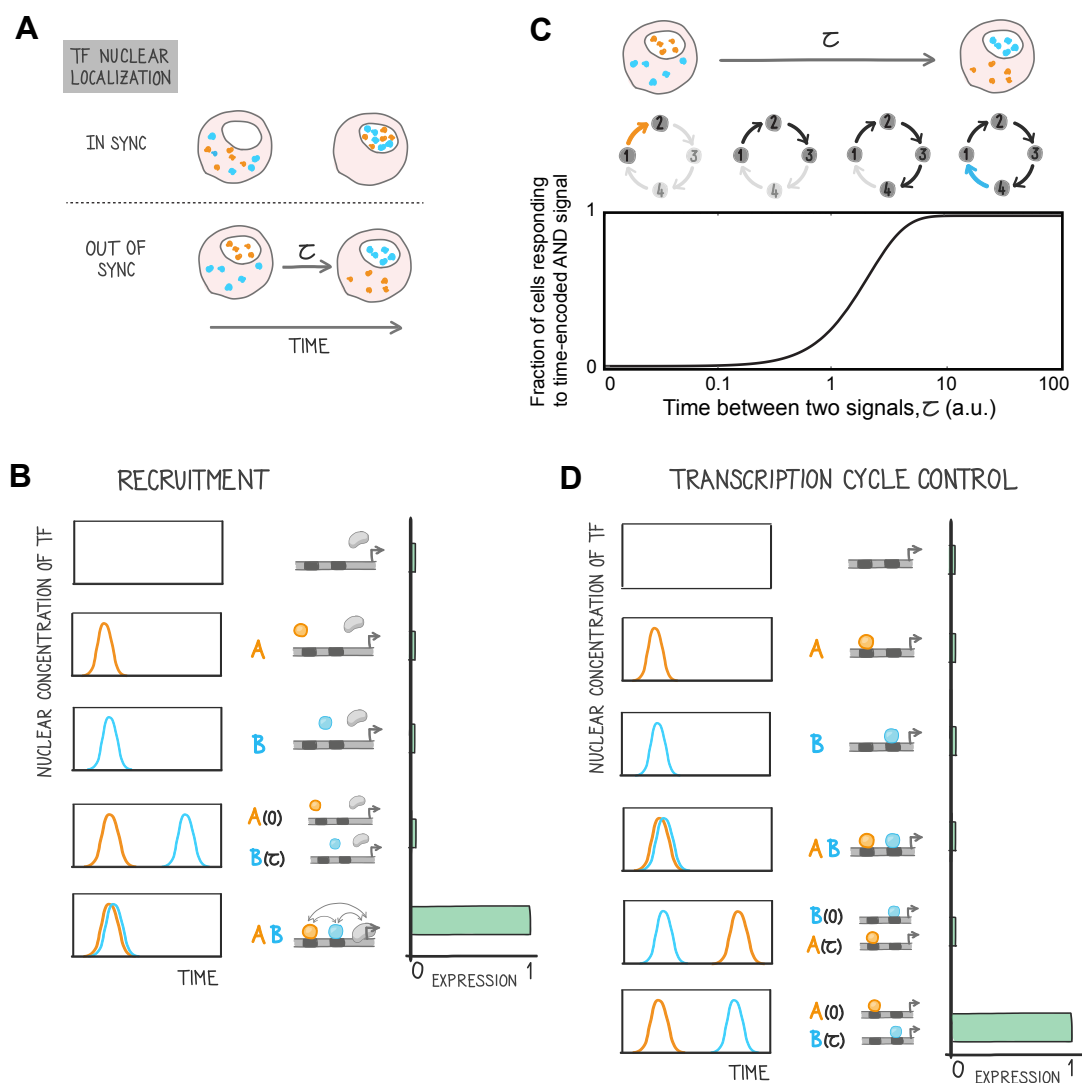


Figure 3. Combinatorial regulation of the transcription cycle enables logical computations on out-of-sync nuclear localization events.

(A) Recent work has shown that transcription can be combinatorially controlled by transient bursts of transcription factor (TF) localization to the nucleus (Lin et al. 2015). Such localization events could be synchronous (upper panel), or asynchronous (lower panel), with a lag between them of time τ , as shown here for two TFs (orange and blue).

(B) Under the “recruitment” model of combinatorial transcriptional regulation, TFs must be simultaneously localized to the nucleus in order for a logical computation such as the AND gate to be possible. This is because cooperative physical interactions between the TFs in binding the DNA and/or recruiting RNA polymerase are required to stabilize the binding of RNAP to the promoter.

(C, D) In contrast, if transcription is combinatorially regulated by TFs acting on the transcription cycle, the TFs do not need to be on the DNA or even in the nucleus at the same time. We examine a transcription cycle with four effective rates (C, middle panel), which allows the implementation of a time encoded AND gate: An initial localization event of TF A (orange) to the nucleus accelerates the first rate, enabling the subsequent middle two steps in the cycle to occur; these two are slow but proceed without regulation from either TF. After a time delay (τ), TF B localizes to the nucleus and enables completion of the last step in the cycle. With increasing time between the bursts of nuclear localization the fraction of cells responding to this temporally-encoded AND gate increases to 1.

(D) Temporal encoding of logical computations by regulating distinct steps in the transcription cycle provides a means for the promoter to distinguish the order in which inputs arrive. For this AND-gate example, a high level of expression cannot be achieved if (i) the time required for steps 1, 2 and 3 to occur is longer than the time that TFs A and B are simultaneously colocalized to the nucleus, or if (ii) B arrives before A.

Null model for combinatorial control of transcription: independent action of TFs

Experimental data are rarely entirely digital. Rather, TFs often have analog effects on the level of transcription (e.g. Sharon et al. 2012; Gertz et al. 2009; Patwardhan et al. 2012). To interpret such analog data, we need a null model of combinatorial regulation. Our goal therefore is to calculate the expected behavior of two activators working independently on the same promoter.

Under the recruitment hypothesis, the independent action of TFs on the bound state of RNAP results in multiplicative activation (Veitia 2003; Bintu et al. 2005b; Herschlag & Johnson 1993). This is because each TF additively lowers the energy of the RNAP-bound state of the promoter; since the energies are contained in the exponent of the expression, the effect on mRNA output is multiplicative (Figure 4A and Supplementary Information).

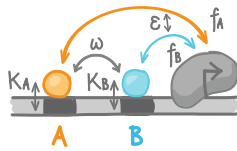
To calculate the effect of independent action of TFs on a transcription cycle, we first consider a simple cycle of only one slow step (Figure 4B). This allows us to compare our results to the recruitment hypothesis, which also implicitly assumes a single rate limiting step. To compare the outcomes of our simulations, we plot the overall rate of transcription in comparison to specific null expectation scenarios (Figure 4). In the “additive” limit, two activators act independently on the same rate limiting step by accelerating it through different pathways; here the expectation is that the transcriptional *enhancements* (Δk_{TXN}) caused by each TF add together, so that $\Delta k_{TXN}(A, B) = \Delta k_{TXN}(A) + \Delta k_{TXN}(B)$. Under the recruitment model, the null expectation is that the *fold-changes* in transcription caused by each TF are multiplicative when both TFs are present. Defining Q as the fold-change in transcription over the basal rate due to the presence of an activator, $Q_{TF} = \frac{k_{TXN}(TF)}{k_{TXN}(0)}$, the null expectation for expression in the recruitment limit when both TFs are present is $Q_{AB} = Q_A \times Q_B$. To visualize how a given mechanism aligns with each null expectation considered above, we construct a null model diagram, where we plot the relative deviation from the “additive” kinetic null model against the relative deviation from the “multiplicative” recruitment null model (Figure 4). Projected on this space, the single-rate recruitment mechanism perfectly aligns with its null, as expected (Figure 4A). The same is true for the single-rate kinetic “additive” mechanism (Figure 4B).

We next consider transcription cycles with two steps, where to formulate a null hypothesis you must choose how TFs will act on the cycle. There are three possible scenarios with two rate limiting steps: TFs accelerate the same step (Figure 4C), different steps (Figure 4D), or both accelerate both steps (Figure 4E). To explore the possible range of outputs one would expect in each case, we randomized the relevant parameters in Equation 5, and projected the results in the null model diagram (Figure 4C-4E). When both TFs work on the same step, the expected distribution is largely sub-additive. The overall transcription rate tends to additivity when the non-regulated step is much faster than the regulated one, such that it has essentially no limiting effect on the rate of transcription. In the second case, when the two TFs work exclusively on different steps, the distribution of rates is always greater than multiplicative. Finally, if both TFs can accelerate both steps, the transcription rate can fall in three of the four quadrants in the null model diagram.

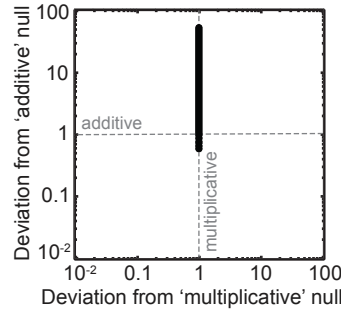
These calculations demonstrate that simple null models of kinetic control over the transcription cycle yield a wide range of possible outputs, depending on the mechanistic assumptions. This emphasizes that, while

ONE STEP

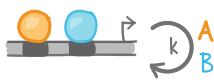
A RECRUITMENT



$$K_{TXN} = \frac{1 + \frac{[A]}{K_A} f_A + \frac{[B]}{K_B} f_B + \frac{[A][B]}{K_A K_B} f_A f_B \omega \varepsilon}{1 + \frac{[A]}{K_A} + \frac{[B]}{K_B} + \frac{[A][B]}{K_A K_B} \omega}$$



B KINETIC



$$K_{TXN} = (1 + a + b) k$$

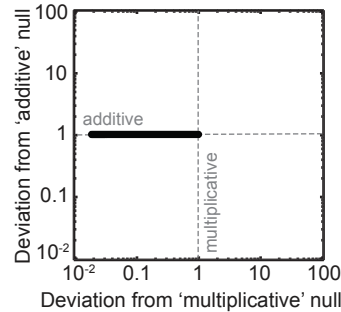


Figure 4. Null models of combinatorial transcriptional control: independent action of activators

To interpret analog data from experiments that examine the effect of transcription factor (TF) number, type and strength on gene expression, we need a null model of combinatorial regulation. Here we define the null model as independent action of TFs. We simulate expression driven by two independently acting activators, using different assumptions about their functions at the promoter. To compare the outcomes, we plot expression relative to the null expectation in two specific limits. In the “recruitment” limit, the null expectation for expression (i.e. expression expected when TFs act independently) is that the fold changes in transcription due to each TF are multiplicative. On the x-axis we therefore plot deviation from this null, $Q_{AB} / (Q_A \times Q_B)$, where Q_{TF} is the fold change in expression over basal. In the “kinetic” limit of two activators working independently on a single slow step, by each working through different pathways, the expectation is that the enhancement of the rate by each TF is additive; on the y-axis we therefore plot deviation from this null expectation, $k_{TXN}(A,B) / (k_{TXN}(A) + k_{TXN}(B))$.

(A) Under the “recruitment” hypothesis expression driven by two independently acting activators is expected to be the product of the fold changes due to each TF acting alone. All parameters in the accompanying equation were randomized, except for ω and ε , which represent interactions between the activators in binding DNA or contacting the promoter and were held at 1. K_A and K_B are the dissociation constants for TFs A and B; we randomized the parameters $\alpha = [A] / K_A$ and $\beta = [B] / K_B$ from 0.01-100. f_A and f_B represent the strength of interaction between each TF and the RNA polymerase and these values were randomized from 1-100.

(B) If either activator can catalyze the rate of the single slow step in transcription via different and independent pathways, the null expectation is that overall enhancement of the rate is the sum of the enhancement due to each TF acting alone. We show this numerically here, where the strength of each activator is randomized between 0-100 while the basal rate, k , is varied between 1-1000.

(C-E) If there is more than one slow rate in the transcription cycle, the null expectation for expression output is more complex and depends on the function of the activators, i.e. which step(s) in the transcription cycle each acts upon. Simulations show the range of possible gene expression outcomes from a simple cycle composed of two slow steps, k_1 and k_2 , controlled by two activators, A and B. The strength of enhancement by A and B of each rate is denoted by a_1, b_1, a_2 and b_2 , and these were varied between 0-100. This is enhancement strength is determined here both by the intrinsic strength of the activator and by its concentration. k_1 and k_2 were randomized between 1-1000.

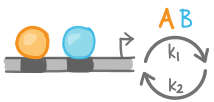
(C) Either activator can catalyze the rate k_1 , independently through different pathways. Their enhancement effects on k_1 sum together (see above), but because of the presence of the other rate, k_2 , their effects on the change in the overall rate of transcription are largely subadditive, tending to additivity in the limit in which $k_2 \gg k_1$.

(D) Two activators working exclusively on different steps in transcription.

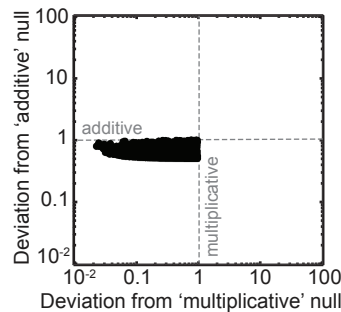
(E) Both activators can work on both steps. For each individual step we assume, as in (B and C), that they catalyze the rate independently through different pathways.

TWO STEPS

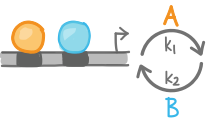
C BOTH ON ONE STEP



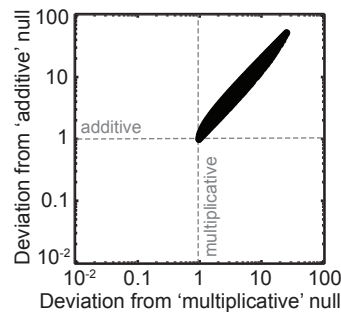
$$K_{TXN} = \frac{(1 + a_1 + b_1) k_1 \times k_2}{(1 + a_1 + b_1) k_1 + k_2}$$



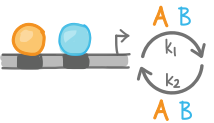
D DIFFERENT STEPS



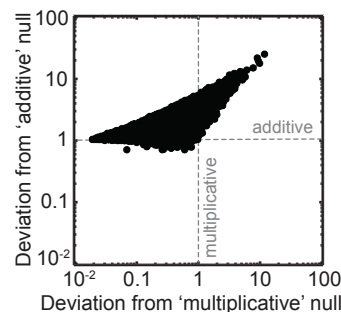
$$K_{TXN} = \frac{(1 + a_1) k_1 \times (1 + b_2) k_2}{(1 + a_1) k_1 + (1 + b_2) k_2}$$



E BOTH ON BOTH STEPS



$$K_{TXN} = \frac{(1 + a_1 + b_1) k_1 \times (1 + a_2 + b_2) k_2}{(1 + a_1 + b_1) k_1 + (1 + a_1 + b_2) k_2}$$



in some cases additivity emerges as a legitimate null model, this is not necessarily so. Therefore, observation of greater- or less-than-additive expression does not necessarily imply that the underlying mechanisms deviate from independent action of TFs.

Quantitatively formulating null hypotheses is one of the major roles for mathematical modeling in biology. Deviations from the predicted null behavior indicates that there is something surprising happening, either a fault in our understanding of the basic biological mechanisms, or an unidentified interaction between the biological components. In the case of transcription, we have defined our null models as two TFs acting independently from each other. Thus, deviations from the null model would imply non-independent action of TFs, which is our definition of synergy. Synergy can arise through a variety of biophysical mechanisms (Box 2); distinguishing these would therefore be a subject for further study.

Box 2: Clarifying the meaning of synergy

The term 'synergy' arises frequently in discussions of how transcription factors (TFs) work together to affect gene expression. Synergy reflects deviations from a null expectation of independent action of components in a complex system (Salvador 2000). In the context of transcription, because the null expectation is not always rigorously defined, synergy is most often used to mean a "greater-than-additive" effect from multiple bound TFs. Although greater-than-additive effects may arise through various mechanisms described below, synergy itself should not be immediately inferred from it because it depends on the null expectation as we show.

There are three main ways that greater-than-additive effects on expression might occur (Herschlag & Johnson 1993; Blau et al. 1996). Notably, two of these reflect null models of independent actions of TFs (and thus cannot actually reflect synergy). This is not an exhaustive list, nor are the mechanisms below mutually exclusive.

(i) Cooperative DNA binding due to direct or indirect recruitment interaction between TFs. In this case greater-than-additive effects should disappear at high concentrations of TFs (Carey et al. 1990; Lin et al. 1990).

(ii) Simultaneous interaction of multiple TFs with the basal transcriptional machinery. If TFs can simultaneously and independently stabilize the basal transcriptional machinery, their free energies will add, leading to a multiplicative effect on the rate of transcription (Veitia 2003; Herschlag & Johnson 1993). For most values of the individual activator contributions, this will be a greater-than-additive effect (it will not be when the fold-change in transcription for each activator is less than 2).

(iii) Kinetic control of different steps of the transcription cycle. If TFs stimulate more than one slow step in transcription then their combined effects will be greater than the sum of their individual actions. This is not the only kinetic means by which greater-than-additive expression could be achieved; for further discussion see Herschlag & Johnson (1993).

BOX: SCIENTIFIC APPLICATION

Predictions arising from combinatorial control of the transcription cycle

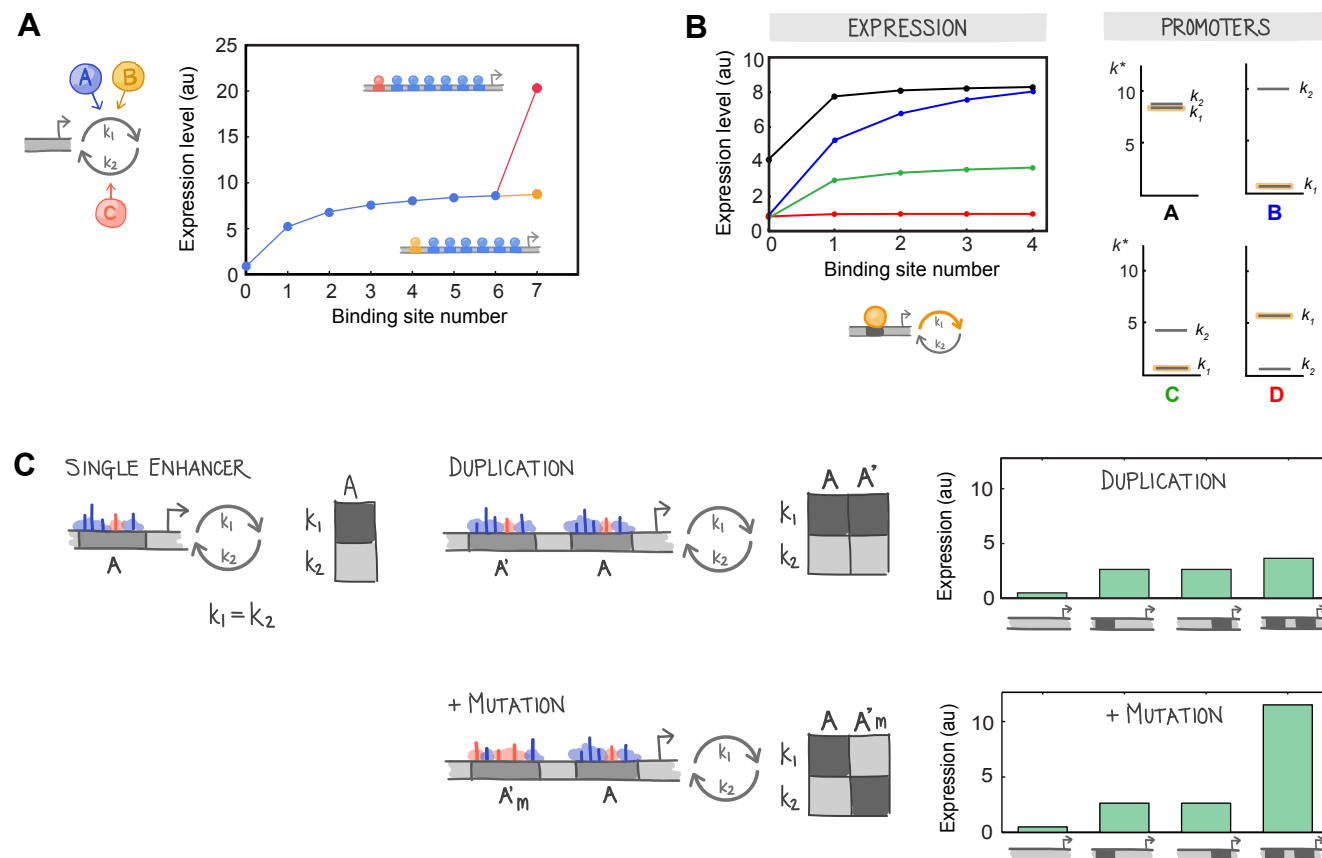
Efforts to uncover a “*cis*-regulatory code” for transcription—a “grammar” that enables prediction of expression from regulatory sequence—have thus far been only moderately successful. Here we discuss three aspects of transcriptional control that might be illuminated by through a kinetic lens. The kinetic view that we have discussed inspires different types of experiments from the focus on transcription factor (TF) binding site arrangement, and makes concrete predictions as to their outcomes.

1. Distinguishing functional classes of TFs

Little structural information exists for TFs outside the DNA binding domain, making it difficult to assess their function roles. Without structural information, we can consider empirical classification of TFs based on their activity. Recent work elucidated functional classes of TFs by assessing their ability to activate transcription in multiple enhancer contexts (Stampfel et al. 2015). Kinetic control also provides a means of functionally classifying TFs. Multimerizing binding sites for an activator eventually saturates transcription (Sharon et al. 2012; Smith et al. 2013; Burz et al. 1998; Carey et al. 1990). If there is only one slow step in transcription (as assumed by “recruitment” models) this expression level reflects the maximum rate of transcription possible from the promoter in question, and no addition of binding sites for any number of TFs will increase the level of expression above this point. However, if there are multiple slow steps, saturating expression reflects having reached the limiting influence of a different rate. Therefore, while adding a TF that is functionally redundant with the first (i.e. acts on the same step) will have no further effect on transcription, adding one that works on a different slow step will boost transcription (see Figure). Limited experiments of this type have been carried out (Blau et al. 1996), but such experiments also readily adapt to high-throughput screening for functional classes of TFs. It would be valuable to quantitatively assay naturally co-occurring transcriptional activators in order to determine how they control different rates in the transcription cycle.

2. Activator-promoter and enhancer-promoter interactions

Core promoter sequence affects the ability of activators or enhancers to activate transcription (Ohler & Wassarman 2010; Butler & Kadonaga 2001; Zabidi et al. 2015), but this is not incorporated into quantitative models of combinatorial transcriptional control. We anticipate that the spectrum of rates in the transcription cycle differs in a meaningful way between core promoters, according to the existence and arrangement of core elements (such as the TATA box, DPE, and Inr). In more extreme cases a promoter could determine whether an activator is able to act on a rate at all, i.e. whether the two are ‘biochemically compatible’ (Juven-Gershon et al. 2008; van Arensbergen et al. 2014). The model we present provides a quantitative framework in which to interpret differences in the expression from enhancer-promoter or activator-promoter pairs. We illustrate this in the accompanying Figure by plotting the expression levels driven by multimerized activator binding sites at four hypothetical core promoters. Because the promoters differ in the relative rates of the two slow steps in this transcription cycle, the effect that any given number of bound activator molecules has on the level of gene expression differs between them. Similarly, the effect of an enhancer bound by activators that work only on the slowest step in transcription would vary widely depending on the ratio between this and next-slowest step. By pairing enhancers with libraries of



Box: Scientific Application. Predictions arising from combinatorial control of the transcription cycle

(A) Distinguishing functional classes of transcription factors (TFs). Different roles of TFs in stimulating the transcription cycle may be distinguished by assessing the effect of a TF on a promoter that is functionally saturated—i.e. where the expression has plateaued. Adding a binding site for second TF that works on the same step in the cycle (orange) will not increase expression, while one that works on a different slow step will do so if that step is now limiting (red). Calculated expression levels are plotted for a two-state promoter, with effective rates $k_1=1$, $k_2=10$ and activator strengths $a_1 = 10$, $b_1 = 2$, $c_2 = 2$.

(B) The role of the promoter in transcriptional control. Multimerizing binding sites for an activator (orange) that works on one step in a two-step cycle produces different expression depending on the relative rates of the two steps at different promoters (right, promoters A-D). Expression calculated for a two-state promoter with effective rates as shown in the plots to the right, and activator strength $a = 10$.

(C) Combinatorial control of transcription by two or more enhancers. Duplicated enhancers may drive little more expression than each enhancer does individually by virtue of responding to the same regulator(s) and thereby accelerating the same rates at the promoter to the same extent. If mutations arise in a duplicate that alter its function, the combined output of the two enhancers may be much higher. Expression calculated for a transcription cycle with two effective rates, where $k_1 = k_2 = 1$ a.u. and the enhancement factors are $a_1 = a'_1 = 10$; $a_2 = a'_2 = 2$; $a'm_1 = 2$, $a'm_2 = 10$.

promoters that have been classified by their core element content (Ohler et al. 2002) it may be possible to work backwards to deduce enhancers' roles in stimulating the transcription cycle.

3. Combinatorial control by two or more enhancers

Many developmental genes contain two or more enhancers that drive spatially and temporally overlapping patterns of expression from the same promoter (Frankel et al. 2010; Hoch et al. 1990; Dunipace et al. 2011; Wunderlich et al. 2015; Staller et al. 2015; Barolo 2011; Prazak et al. 2010; Perry et al. 2011). The mechanism by which the information from these “shadow” enhancers is combined in single cells to drive a given rate of transcription is unknown and the combined output is unpredictable (Dunipace et al. 2011; Bothma et al. 2015; Mas et al. 2004). It is plausible that shadow enhancers work together by regulating different steps in the transcription cycle. The strength of effect that an enhancer has on each rate in the transcription cycle will depend on the number and type of TFs bound to it in any given cell. But since TF concentrations change over space and time, so too may the the enhancer's effect on the transcription cycle and how the expression it drives is integrated with that from other enhancers. The common experiment to test for interactions between enhancers (measuring the influence of each alone and together) is difficult to rigorously interpret since multiple models could explain the experimental data (see Figure 4). A variation on this experiment that may prove useful is to compare the effects of enhancer duplications to those of shadow enhancers. We speculate that a pair of shadow enhancers may behave differently than a duplication of either. Duplicate enhancers contain an identical number and type of transcriptional activators, which makes them likely to work in the same manner on the transcription cycle, leading to a combined expression output that is largely subadditive. If mutations accumulate in a duplicated pair of enhancers their functions could diverge, leading to different effects on transcription. This is analogous to some sibling enhancer pairs, which are known to respond to different activators (Wunderlich et al. 2015; Staller et al. 2015).

DISCUSSION

We argue that a quantitative model of combinatorial gene regulation should include multiple steps in the transcription cycle. We show that two TFs acting independently can produce a wide range of outputs, both analog and digital. These outputs depend on the underlying rates of the transcription cycle and how the TFs act on these rates. This view of transcriptional control emphasizes the information-processing role of the promoter, which dictates the spectrum of effective rates in the transcription cycle. We speculate that combinatorial control of the transcription cycle enables greater plasticity than a process that relies only on physical interactions between TFs; this could enable rapid regulatory evolution as is observed in many systems.

A cis-regulatory “code” at the level of TF function?

The holy grail of deciphering regulatory DNA is unlocking the “cis-regulatory code”, which ideally would allow us to predict gene expression driven by any regulatory sequence. The emphasis in efforts to crack this code has been on the number, affinity and arrangement of TF binding sites in enhancers (reviewed in Weingarten-Gabbay & Segal 2014). This emphasis is reasonable if combinatorial control of transcription is channeled through regulation of a single slow step in the transcription cycle. However, while binding site arrangement does play an important role in cooperative DNA binding (Thanos & Maniatis 1995; Adams & Workman 1995), a general code at the level of TF binding sites has remained elusive (Weingarten-Gabbay & Segal 2014). We have shown here that kinetic control of the transcription cycle alone can also produce complex computations. From this kinetic perspective, it may matter less how TF binding sites are configured and matter more which TFs are recruited to regulatory DNA. Recent work categorizes TFs into functional classes based on their combinatorial activity in different enhancer contexts (Stampfel et al. 2015), which offers the opportunity to test the kinetic perspective directly.

Kinetic control of the transcription cycle casts the basal promoter in a key role. While some communities recognize the information processing role of the promoter (Ohler & Wassarman 2010; Danino et al. 2015; Zehavi et al. 2014), others consider enhancers the source of spatiotemporal patterns of expression, while basal promoters are a passive elements that make genes competent for transcription to a greater or lesser extent (“strong” versus “weak” promoters). We show that TF function is necessarily interpreted through the rate-limiting steps of the promoter, firmly putting the promoter in an information processing role in a quantitative modeling framework. This framework can now accommodate multiple experimental observations. For example, the composition of core motifs in basal promoters plays a key role in cell type-specific gene regulation (reviewed in (Danino et al. 2015; Zehavi et al. 2014; Ohler & Wassarman 2010)), affecting the composition of the preinitiation complex (Lewis et al. 2005), pausing (Amir-Zilberstein et al. 2007; Hendrix et al. 2008), and termination and recycling of RNAP (Mapendano et al. 2010). Moreover, activation domains have different abilities to work with different core promoter elements, emphasizing how these two types of regulatory DNA work together (Emami et al. 1995; Das et al. 1995; Juven-Gershon et al. 2008).

The model of transcription cycle control readily accommodates endogenous complexities

Building models necessarily involves deciding which features to include and which to leave out. Our goal in this paper has been to highlight how combinatorial regulation of the transcription cycle may be exploited by cells to process information, an area that has received relatively little attention. To accomplish

this we purposely simplified the mechanisms by which TFs bind, unbind and interact with the transcriptional machinery. We used a mean field approximation, where the effect of TFs on individual rates were described by their "activity", which corresponds to the product of the rate of activation when a TF is bound and the probability that the TF is indeed bound. The latter is a function of its concentration, binding affinity and mode of binding (independent, cooperative, etc.). In practice, activation by a TF is a single molecule stochastic process and thus our approach is only an approximation that may very well not capture the actual biophysics in many instances. However, while these details are important on a case-by-case basis, their abstraction into a single effective parameter is what allowed us to examine the transcription cycle in isolation. We note that a full model incorporating both the stochastic TF association/dissociation and the mechanistic step of accelerating the rates that they regulate can be easily accomplished using the same formalism that we employed here.

For the purposes of illustration we considered a very simple scenario of combinatorial control of the transcription cycle consisting of one or two slow steps. Of course, transcription is much more complex. For example, promoters differ in which and how many rates are limiting. The model we describe is agnostic to the identities of the slow steps in a transcription cycle. What determines the overall rate of transcription is the relative rates and how TFs impinge upon them, both of which can be computationally explored. There is also not a one-to-one mapping of TF to function, i.e. a single TF can work on multiple steps either directly or via the cofactors with which it interacts (reviewed in (Fuda et al. 2009)). For example, the yeast activator Gcn4 contacts four coregulator complexes, each of which affect chromatin remodeling, assembly of the preinitiation complex and elongation to different degrees (Fishburn et al. 2005; Govind et al. 2005). Reciprocally, a single process can be influenced by multiple TFs, often through shared intermediates like Mediator. In our model, the catalysis of a given rate can be arbitrarily apportioned between TFs, and vice versa. Finally, the concentrations and cooperativities in binding to DNA will also play roles in determining the outcome of transcription. Equilibrium models based on statistical thermodynamics have already explored these in detail (Bintu, Buchler, Garcia, Gerland, Hwa, Kondev, Kuhlman, et al. 2005; Bintu, Buchler, Garcia, Gerland, Hwa, Kondev & Phillips 2005). The purpose of our analysis has been to complement these models with the kinetic aspects that are most apparent when the binding sites are already saturated.

Control of the transcription cycle may enable evolutionary plasticity

Here we have addressed how a promoter could execute computations without relying on interactions between transcription factors. But an interesting question going forwards will be what the advantages and disadvantages are of generating computations through TF interactions, versus by controlling different kinetic steps in transcription, with respect to how regulatory evolution can occur. We speculate that control of the transcription cycle could help to enable the flexibility apparent in how gene expression patterns are encoded in regulatory DNA between species (Ludwig et al. 2005; Wunderlich et al. 2015; Tuch et al. 2008; Tsong et al. 2006). As we have discussed, by combinatorially controlling the transcription cycle, the the same computations can be encoded without the need for any cooperation between transcription factors. This relieves the need for interacting TF binding sites to be in close vicinity with one another, removing a significant constraint on how binding sites are organised within regulatory sequences. In addition, we suggest that kinetic control might also enable changes to accumulate in regulatory DNA through "buffering" their potential phenotypic consequences. Binding sites for TFs that work on fast rates may come and go without affecting gene regulation, but such variation could be unmasked by mutations in

promoter sequence, or environmental or physiological changes. This is analogous to the evolutionary capacitor function of molecular chaperones (Rutherford & Lindquist 1998).

Outlook

Given the overwhelming evidence for action of TFs on multiple distinct, slow steps in transcription, it is time to incorporate this reality into models of combinatorial gene regulation. Viewing transcriptional regulation through a kinetic lens may help to unify and explain multiple experimental realities, such as complex transcription cycles, fast TF binding, and flexible regulatory evolution. In the long term, elucidating TF function beyond simple ‘activation’ or ‘repression’ may finally reveal a tractable *cis*-regulatory code.

EXPERIMENTAL PROCEDURES

Calculating mean mRNA production from an arbitrarily complex promoter

Due to the stochastic nature of transcription, the instantaneous rate of initiation will fluctuate over time, even under a constant environment. To calculate the moments of the distributions of transcription rates within a single cell over time (which, by assuming ergodicity also reflect the moments of the distribution of initiation rates within a population of genetically homogeneous cells exposed to the same environment), we use a matrix formalism akin to that which we introduced in previous work (Sanchez & Kondev 2008; Sanchez et al. 2011; Sanchez et al. 2013), inspired by earlier models of a two-state promoter (Peccoud & Ycart 1995; Kepler & Elston 2001).

To illustrate this formalism, consider a promoter that can exist in three states. Transitions rates from state i to j is described by k_{ij} . We assume in this example that the only transitions allowed are: from 1 and 2 (in both directions, described by k_{12} and k_{21}), from 2 and 3 (in both directions: k_{23} and k_{32}), and finally that a productive transition (i.e. one that is accompanied by the synthesis of a new mRNA molecule) is allowed from 3 to 1 (given by k_{31}); this last transition is irreversible. We introduce a vector containing the steady state probabilities of finding the promoter in states 1, 2 and 3: $p = (p_1, p_2, p_3)$. The average rate of transcription will be equal to the dot product of the vector of mRNA synthesis from each state $r = (0, 0, k_{31})$ and the probability vector p : $K_{TXN} = r \cdot p = k_{31}p_3$. In order to calculate how p depends on the kinetic rates, we just need to solve the equation: $\hat{K} \cdot p = u$, where u is the vector of the same dimension as p (which we’ll define as L), and whose entries are $u_j = \delta_{jL}$. In turn, \hat{K} is the matrix with entries $K_{ij} = \delta_{iL} + (1 - \delta_{iL}) \left((1 - \delta_{ij})k_{ij} - \delta_{ij} \sum_j k_{ij} \right)$. In both of these equations δ_{ij} represents the Kronecker delta, i.e. $\delta_{ij} = 1$ if $i=j$, and $\delta_{ij} = 0$ otherwise.

Combinatorial regulation of transcription by bursts of nuclear localization.

The fraction of cells that have produced a transcript in response to the nuclear localization of A and B for a four step kinetic cycle such as that depicted in Figure 3 can be determined as follows. First, we define the

promoter state vector $\mathbf{p}(t) = (p_1(t), p_2(t), p_3(t), p_4(t))$, which contains the probabilities that a promoter will be at each state (or what is the same, the fraction of cells in a population that are at those states) at time t . To calculate the fraction of cells that have produced at least one transcript in response to the nuclear localization events, we consider the transition from state 4 to state 1 (which produces a transcript) as leading to an absorbing boundary (this is a convenient mathematical artifice with no bearing on the actual biochemistry). The value of the vector $\mathbf{p}(t)$ can be found by solving the equation: $\frac{d}{dt}\mathbf{p} = \hat{K} \cdot \mathbf{p}$, where \hat{K} in this case has elements: $K_{11} = -K_{21} = -k_{12}$, $K_{22} = -K_{32} = -k_{23}$, $K_{33} = -K_{43} = -k_{34}$ and $K_{44} = -K_{14} = -k_{41}$. For all other values of i, j , $K_{ij} = 0$. In the example we worked out in the main text, we assumed that the first step in the cycle, k_{12} , depends on the intranuclear concentration of A, whereas the final step in the cycle, k_{41} , depends on the intranuclear concentration of B. Since both $[A]$ and $[B]$ change with time as a result of a nuclear localization burst, this leads to time dependent functions which we write as: $k_{12} = f_A(t)$, $k_{41} = f_B(t)$. In order to investigate an example of this model that can be solved analytically, in simple terms but without losing generality, we assume that $f_A(t)$ is such that it causes essentially all cells to complete the first step of the cycle in a very short time (in comparison to the time required to complete the second step of the cycle). In practice, this can be modeled by a square step function with very large height (k_A) and very narrow width (T), so that $k_A T \gg 1 \gg k_{23} T$. Likewise, we also assume that $f_B(t)$ has the same shape as $f_A(t)$, and we also assume the transitions from state 2 to state 3, and from state 3 to state 4 both happen with the same rate k_t , which is independent of the intranuclear concentrations of A and B. Solving this system of equations under the constraints and assumptions discussed above yields Equation 6 in the text

Simulating analog and digital computations through control of the transcription cycle

To explore the space of mean transcription rates that could be achieved under a given model of activator function on a simple two-step cycle (Figure 4), we randomized key parameters in Equation 5 in the main text. Specifically, we randomized the rates of the two slow steps in the cycle over three orders of magnitude and varied activator strength, described by $\{a_1, a_2, b_1, b_2\}$, between 0 and 100; this is within the same order of magnitude as the measured effect of an activator on a given rate in transcription (Friedman & Gelles 2012). To explore the transcription rates that can be produced under the recruitment model, we randomized the parameters in the equation shown in Figure 4A (Bintu et al. 2005a). Because we were interested only in the null situation of activator independence, we held ω and ϵ at 1 (these describe cooperativity between TFs in binding to DNA and in contacting RNAP, respectively). The concentrations of activators A and B ($[A]$ and $[B]$) were varied between 0.01 - 100, while f_A and f_B were varied between 1 - 100.

To investigate potential means of achieving logical calculations from a cycle of two slow effective rates controlled by two transcription factors we surveyed a range of parameters and chose a set that worked well for illustration. The parameter values used are listed in Table S1.

AUTHOR CONTRIBUTIONS

C.S., A.H.D and A.S. conceived and designed the study and wrote the manuscript. C.S. and A.S. developed the analytical framework.

ACKNOWLEDGEMENTS

We would like to thank Hernan Garcia, Jané Kondev and Sandeep Choubey for many useful discussions, and Stirling Churchman, Jeremy Gunawardena, Mike Springer, Johan Paulsson and Jané Kondev for their comments and feedback on the manuscript. A.S. was supported by a Rowland Junior Fellowship from The Rowland Institute at Harvard. C.S. was supported by a Harvard Graduate Society Research Fellowship. Work in the DePace lab is supported by NIH U01 GM103804-01A1, NSF CAREER IOS-1452557, and the Grinnell Fund for Biomedical Research.

REFERENCES

- Adams, C.C. & Workman, J.L., 1995. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Molecular and cellular biology*, 15(3), pp.1405–1421.
- Adelman, K. & Lis, J.T., 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature reviews. Genetics*, 13(10), pp.720–731.
- Amir-Zilberstein, L. et al., 2007. Differential Regulation of NF- κ B by Elongation Factors Is Determined by Core Promoter Type. *Molecular and cellular biology*, 27(14), pp.5246–5259.
- van Arensbergen, J., van Steensel, B. & Bussemaker, H.J., 2014. In search of the determinants of enhancer–promoter interaction specificity. *Trends in cell biology*, 24(11), pp.695–702.
- Barboric, M. & Peterlin, B.M., 2005. A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. *PLoS biology*, 3(2), p.e76.
- Barolo, S., 2011. Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 34(2), pp.135–141.
- Batchelor, E. et al., 2011. Stimulus-dependent dynamics of p53 in single cells. *Molecular systems biology*, 7, p.488.
- Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., et al., 2005b. Transcriptional regulation by the numbers: applications. *Current opinion in genetics & development*, 15(2), pp.125–135.
- Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J. & Phillips, R., 2005a. Transcriptional regulation by the numbers: models. *Current opinion in genetics & development*, 15(2), pp.116–124.
- Blair, W.S., Fridell, R.A. & Cullen, B.R., 1996. Synergistic enhancement of both initiation and elongation by acidic transcription activation domains. *The EMBO journal*, 15(7), pp.1658–1665.
- Blau, J. et al., 1996. Three functional classes of transcriptional activation domain. *Molecular and cellular biology*, 16(5), pp.2044–2055.
- Bothma, J.P. et al., 2015. Enhancer additivity and non-additivity are determined by enhancer strength in the *Drosophila* embryo. *eLife*, 4. Available at: <http://dx.doi.org/10.7554/eLife.07956>.
- Brewster, R.C., Jones, D.L. & Phillips, R., 2012. Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. *PLoS computational biology*, 8(12), p.e1002811.
- Browning, D.F. & Busby, S.J., 2004. The regulation of bacterial transcription initiation. *Nature reviews. Microbiology*, 2(1), pp.57–65.
- Brown, S.A. et al., 1998. Transcriptional activation domains stimulate initiation and elongation at different times and via different residues. *The EMBO journal*, 17(11), pp.3146–3154.
- Buchler, N.E., Gerland, U. & Hwa, T., 2003. On schemes of combinatorial transcription logic. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9),

- pp.5136–5141.
- Burz, D.S. et al., 1998. Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the *Drosophila* embryo. *The EMBO journal*, 17(20), pp.5998–6009.
- Butler, J.E. & Kadonaga, J.T., 2001. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes & development*, 15(19), pp.2515–2519.
- Cai, L., Dalal, C.K. & Elowitz, M.B., 2008. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*, 455(7212), pp.485–490.
- Carey, M. et al., 1990. A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature*, 345(6273), pp.361–364.
- Carey, M., 1998. The enhanceosome and transcriptional synergy. *Cell*, 92(1), pp.5–8.
- Choubey, S., Kondev, J. & Sanchez, A., 2015. Deciphering Transcriptional Dynamics In Vivo by Counting Nascent RNA Molecules. *PLoS computational biology*, 11(11), p.e1004345.
- Coulon, A. et al., 2013. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature reviews. Genetics*, 14(8), pp.572–584.
- Dalal, C.K. et al., 2014. Pulsatile dynamics in the yeast proteome. *Current biology: CB*, 24(18), pp.2189–2194.
- Danino, Y.M. et al., 2015. The core promoter: At the heart of gene expression. *Biochimica et biophysica acta*, 1849(8), pp.1116–1131.
- Das, G., Hinkley, C.S. & Herr, W., 1995. Basal promoter elements as a selective determinant of transcriptional activator function. *Nature*, 374(6523), pp.657–660.
- Davidson, E.H., 2011. Evolutionary bioscience as regulatory systems biology. *Developmental biology*, 357(1), pp.35–40.
- Dresch, J.M., Richards, M. & Ay, A., 2013. A primer on thermodynamic-based models for deciphering transcriptional regulatory logic. *Biochimica et biophysica acta*, 1829(9), pp.946–953.
- Dunipace, L., Ozdemir, A. & Stathopoulos, A., 2011. Complex interactions between cis-regulatory modules in native conformation are critical for *Drosophila* snail expression. *Development*, 138(18), pp.4075–4084.
- Emami, K.H., Navarre, W.W. & Smale, S.T., 1995. Core promoter specificities of the Sp1 and VP16 transcriptional activation domains. *Molecular and cellular biology*, 15(11), pp.5906–5916.
- Fishburn, J., Mohibullah, N. & Hahn, S., 2005. Function of a eukaryotic transcription activator during the transcription cycle. *Molecular cell*, 18(3), pp.369–378.
- Frankel, N. et al., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*, 466(7305), pp.490–493.
- Frank, T.D., Carmody, A.M. & Kholodenko, B.N., 2012. Versatility of Cooperative Transcriptional

- Activation: A Thermodynamical Modeling Analysis for Greater-Than-Additive and Less-Than-Additive Effects M. Isalan, ed. *PLoS one*, 7(4), p.e34439.
- Friedman, L.J. & Gelles, J., 2012. Mechanism of Transcription Initiation at an Activator-Dependent Promoter Defined by Single-Molecule Observation. *Cell*, 148(4), pp.679–689.
- Fuda, N.J., Ardehali, M.B. & Lis, J.T., 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*, 461(7261), pp.186–192.
- Garcia, H.G. & Phillips, R., 2011. Quantitative dissection of the simple repression input–output function. *Proceedings of the National Academy of Sciences*, 108(29), pp.12173–12178.
- Gertz, J., Siggia, E.D. & Cohen, B.A., 2009. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, 457(7226), pp.215–218.
- Gonzalez-Couto, E., Klages, N. & Strubin, M., 1997. Synergistic and promoter-selective activation of transcription by recruitment of transcription factors TFIID and TFIIB. *Proceedings of the National Academy of Sciences of the United States of America*, 94(15), pp.8036–8041.
- Govind, C.K. et al., 2005. Simultaneous recruitment of coactivators by Gcn4p stimulates multiple steps of transcription in vivo. *Molecular and cellular biology*, 25(13), pp.5626–5638.
- Greber, D. & Fussenegger, M., 2007. Mammalian synthetic biology: engineering of sophisticated gene networks. *Journal of biotechnology*, 130(4), pp.329–345.
- Green, M.R., 2005. Eukaryotic transcription activation: right on target. *Molecular cell*, 18(4), pp.399–402.
- Gross, C.A. et al., 1998. The functional and regulatory roles of sigma factors in transcription. In *Cold Spring Harbor symposia on quantitative biology*. Cold Spring Harbor Laboratory Press, pp. 141–156.
- Gunawardena, J., 2014. Models in biology: “accurate descriptions of our pathetic thinking.” *BMC biology*, 12, p.29.
- Hahn, S., 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nature structural & molecular biology*, 11(5), pp.394–403.
- Hao, N. & O’Shea, E.K., 2012. Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nature structural & molecular biology*, 19(1), pp.31–39.
- Hendrix, D.A. et al., 2008. Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America*, 105(22), pp.7762–7767.
- Herschlag, D. & Johnson, F.B., 1993. Synergism in transcriptional activation: a kinetic view. *Genes & development*, 7(2), pp.173–179.
- He, X. et al., 2010. Thermodynamics-Based Models of Transcriptional Regulation by Enhancers: The Roles of Synergistic Activation, Cooperative Binding and Short-Range Repression U. Ohler, ed. *PLoS computational biology*, 6(9), p.e1000935.
- Hoch, M. et al., 1990. cis-acting control elements for Krüppel expression in the *Drosophila* embryo. *The*

- EMBO journal*, 9(8), pp.2587–2595.
- Istrail, S. & Davidson, E.H., 2005. Logic functions of the genomic cis-regulatory code. *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp.4954–4959.
- Jonkers, I., Kwak, H. & Lis, J.T., 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife*, 3, p.e02407.
- Juven-Gershon, T., Hsu, J.-Y. & Kadonaga, J.T., 2008. Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. *Genes & development*, 22(20), pp.2823–2830.
- Kepler, T.B. & Elston, T.C., 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophysical journal*, 81(6), pp.3116–3136.
- Keung, A.J. et al., 2014. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell*, 158(1), pp.110–120.
- Lahav, G. et al., 2004. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature genetics*, 36(2), pp.147–150.
- Lewis, B.A. et al., 2005. Functional characterization of core promoter elements: DPE-specific transcription requires the protein kinase CK2 and the PC4 coactivator. *Molecular cell*. Available at: <http://www.sciencedirect.com/science/article/pii/S1097276505012499>.
- Lin, Y. et al., 2015. Combinatorial gene regulation by modulation of relative pulse timing. *Nature*, 527(7576), pp.54–58.
- Lin, Y.S. et al., 1990. How different eukaryotic transcriptional activators can cooperate promiscuously. *Nature*, 345(6273), pp.359–361.
- Liu, M. et al., 2003. Kinetics of transcription initiation at lacP1. Multiple roles of cyclic AMP receptor protein. *The Journal of biological chemistry*, 278(41), pp.39755–39761.
- Ludwig, M.Z. et al., 2005. Functional evolution of a cis-regulatory module. *PLoS biology*, 3(4), p.e93.
- Mapendano, C.K. et al., 2010. Crosstalk between mRNA 3' end processing and transcription initiation. *Molecular cell*, 40(3), pp.410–422.
- Mas, J.-A., García-Zaragoza, E. & Cervera, M., 2004. Two functionally identical modular enhancers in *Drosophila* troponin T gene establish the correct protein levels in different muscle types. *Molecular biology of the cell*, 15(4), pp.1931–1945.
- McClure, W.R., 1985. Mechanism and control of transcription initiation in prokaryotes. *Annual review of biochemistry*, 54, pp.171–204.
- Nechaev, S. & Adelman, K., 2011. Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochimica et biophysica acta*, 1809(1), pp.34–45.
- Ohler, U. et al., 2002. Computational analysis of core promoters in the *Drosophila* genome. *Genome biology*, 3(12), p.RESEARCH0087.
- Ohler, U. & Wassarman, D.A., 2010. Promoting developmental transcription. *Development*, 137(1),

pp.15–26.

- Patwardhan, R.P. et al., 2012. Massively parallel functional dissection of mammalian enhancers in vivo. *Nature biotechnology*, 30(3), pp.265–270.
- Peccoud, J. & Ycart, B., 1995. Markovian Modeling of Gene-Product Synthesis. *Theoretical population biology*, 48(2), pp.222–234.
- Perry, M.W., Boettiger, A.N. & Levine, M., 2011. Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America*, 108(33), pp.13570–13575.
- Phillips, R., 2015. Theory in Biology: Figure 1 or Figure 7? *Trends in cell biology*, 25(12), pp.723–729.
- Prazak, L., Fujioka, M. & Gergen, J.P., 2010. Non-additive interactions involving two distinct elements mediate sloppy-paired regulation by pair-rule transcription factors. *Developmental biology*, 344(2), pp.1048–1059.
- Ptashne, M. & Gann, A., 1997. Transcriptional activation by recruitment. *Nature*, 386(6625), pp.569–577.
- Reppas, N.B. et al., 2006. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Molecular cell*, 24(5), pp.747–757.
- Rutherford, S.L. & Lindquist, S., 1998. Hsp90 as a capacitor for morphological evolution. *Nature*, 396(6709), pp.336–342.
- Salvador, A., 2000. Synergism analysis of biochemical systems. I. Conceptual framework. *Mathematical biosciences*, 163(2), pp.105–129.
- Sanchez, A. et al., 2011. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS computational biology*, 7(3), p.e1001100.
- Sanchez, A., Choubey, S. & Kondev, J., 2013. Regulation of noise in gene expression. *Annual review of biophysics*, 42, pp.469–491.
- Sanchez, A. & Kondev, J., 2008. Transcriptional control of noise in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), pp.5081–5086.
- Sharon, E. et al., 2012. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nature biotechnology*, 30(6), pp.521–530.
- Sherman, M.S. & Cohen, B.A., 2012. Thermodynamic state ensemble models of cis-regulation. *PLoS computational biology*, 8(3), p.e1002407.
- Smith, R.P. et al., 2013. Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *Nature genetics*, 45(9), pp.1021–1028.
- Spitz, F. & Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics*, pp.1–14.
- Staller, M.V. et al., 2015. Shadow enhancers enable Hunchback bifunctionality in the *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America*, 112(3),

pp.785–790.

- Stampfel, G. et al., 2015. Transcriptional regulators form diverse groups with context-dependent regulatory functions. *Nature*. Available at: <http://dx.doi.org/10.1038/nature15545> [Accessed November 9, 2015].
- Stasevich, T.J. et al., 2014. Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature*, 516(7530), pp.272–275.
- Struhl, K., 1991. Mechanisms for diversity in gene expression patterns. *Neuron*, 7(2), pp.177–181.
- Thanos, D. & Maniatis, T., 1995. Virus induction of human IFN β gene expression requires the assembly of an enhanceosome. *Cell*, 83(7), pp.1091–1100.
- Tsong, A.E. et al., 2006. Evolution of alternative transcriptional circuits with identical logic. *Nature*, 443(7110), pp.415–420.
- Tuch, B.B., Li, H. & Johnson, A.D., 2008. Evolution of eukaryotic transcription circuits. *Science*, 319(5871), pp.1797–1799.
- Veitia, R.A., 2003. A sigmoidal transcriptional response: cooperativity, synergy and dosage effects. *Biological reviews of the Cambridge Philosophical Society*, 78(1), pp.149–170.
- Wade, J.T. & Struhl, K., 2008. The transition from transcriptional initiation to elongation. *Current opinion in genetics & development*, 18(2), pp.130–136.
- Wang, J. et al., 1999. A mathematical model for synergistic eukaryotic gene activation. *Journal of molecular biology*, 286(2), pp.315–325.
- Weingarten-Gabbay, S. & Segal, E., 2014. The grammar of transcriptional regulation. *Human genetics*, 133(6), pp.701–711.
- Wunderlich, Z. et al., 2015. Krüppel Expression Levels Are Maintained through Compensatory Evolution of Shadow Enhancers. *Cell reports*, 12(11), pp.1740–1747.
- Yankulov, K. et al., 1994. Transcriptional elongation by RNA polymerase II is stimulated by transactivators. *Cell*, 77(5), pp.749–759.
- Zabidi, M.A. et al., 2015. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature*, 518(7540), pp.556–559.
- Zehavi, Y. et al., 2014. The core promoter composition establishes a new dimension in developmental gene networks. *Nucleus*, 5(4), pp.298–303.
- Zeitlinger, J. et al., 2007. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature genetics*, 39(12), pp.1512–1516.
- Zoller, B. et al., 2015. Structure of silent transcription intervals and noise characteristics of mammalian genes. *Molecular systems biology*, 11(7), p.823.

FIGURE LEGENDS

Figure 1. Promoters integrate information from multiple regulatory sequences

(A) Multiple regulatory sequences (top) or individual bound transcriptional activators (bottom) impinge on a given promoter at the same time. Together these regulatory sequences determine the level and timing of gene expression. We ask how these inputs are integrated by the promoter, by defining the overall expression level driven by two elements (either enhancers or activators) as a function of their individual effects.

(B) Two regulatory elements, A and B, may combine in different ways to determine the overall expression level.

(C) Models of combinatorial transcriptional control usually assume that the single limiting step is “recruitment” of RNA Polymerase (RNAP) to the promoter. Recruitment is mediated by physical interactions between different transcription factors (TFs, A and B) and the DNA, and between TFs and RNAP. The rate of transcription is assumed to be proportional to RNAP occupancy of the promoter. K_A – dissociation constant for TF A. f_A – strength of interaction between TF A and RNAP. ω – cooperativity between TFs A and B in binding to DNA. ε – cooperativity between the TFs in recruiting RNAP.

(D) Recruitment models implicitly assume a single slow effective rate in transcription. Transcription is limited by the dissociation of RNAP from the promoter, and activators work by reducing this dissociation rate. The single effective rate is described by the rate constant k (first column). Adding an activator (orange) accelerates this effective rate, raising the overall rate of transcription (k_{TXN} , arrow). All other steps in transcription are assumed to be infinitely fast (not shown).

(E) The transcription cycle contains numerous distinct steps (Box 1). Shown here are the effective rates of N arbitrary steps, with shorter arrows indicating faster rates. TFs may work on the same or different steps in the cycle. Here, TF A accelerates effective rate k_i (orange) while B accelerates one of the faster, unlabeled rates (blue).

(F) The effective rates of the transcription cycle are depicted before and after addition of an activator (orange) that works on the slowest step. The overall rate of transcription increases when the activator is added, but is limited by the next slowest rate.

Box 1. The transcription cycle in prokaryotes and eukaryotes

(A) Prokaryotes. RNAP is recruited to the promoter through association with a sigma factor to form the closed complex (Gross et al. 1998; Browning & Busby 2004). Subsequent melting of the DNA forms the open complex (McClure 1985) and RNAP then escapes the promoter into elongation. It was widely assumed that recruitment was the rate-limiting step in bacterial transcription (Ptashne & Gann 1997); however, the transition from initiation to elongation is rate limiting at many promoters (Reppas et al. 2006; Wade & Struhl 2008). Diagram of a bacterial transcription cycle adapted from Friedman & Gelles (2012).

(B) We used the forward and reverse reaction rates measured in Friedman & Gelles (2012) to calculate the “effective” forward rates through the transcription cycle (Results and Figure S1A-C). As this rates diagram shows, assuming a single slow step in transcription may be problematic even for bacterial promoters: the two slowest effective rates are within an order of magnitude of one another.

(C) Eukaryotes. The transcription cycle begins with the formation of the preinitiation complex comprised of RNAP and general TFs (Fuda et al. 2009). The chromatin surrounding the promoter is modified, which is coordinated by DNA-bound TFs that recruit chromatin remodelers and nucleosome modifying enzymes. Transcription initiation can take place once the DNA has been unwound to form the open complex. RNAP

must subsequently escape the grip of proteins at the promoter to enter into elongation. At most genes in higher eukaryotes, RNAP pauses downstream of the promoter and must be actively released from this paused state into productive elongation through the body of the gene (reviewed in (Zeitlinger et al. 2007; Jonkers et al. 2014; Adelman & Lis 2012)). Diagram of higher eukaryotic transcription cycle adapted from Fuda et al. (2009) and Boettiger 2013.

(D) The “effective” rates for a simplified eukaryotic transcription cycle were calculated from Stasevich et al (2014) (Results and Figure S1D-F).

Figure 2. Logical computations could be explained by regulation of the transcription cycle

(A) Summary table showing how combinations of TFs A and B map to common Boolean or logical computations. The input TFs are either at low or high concentration, and the output is binary: the gene is ON (highly expressed) or OFF (very lowly expressed, or not expressed at all).

(B) We consider a simple two-step transcription cycle. The effective forward rates are characterized by the rate constants k_1 and k_2 , and a molecule of mRNA is produced during step two.

(C-H) Control of a simple two-step transcription cycle by two TFs can produce all of the basic Boolean computations. For each logic gate we show an example of how TFs A and B could bring about the computation by acting on one or both rates (grey horizontal bars), either as activators (arrows) or repressors (blunted arrows). The resulting expression output due to each TF acting individually or in combination is shown in green, normalised to the maximum expression in each plot such that 0 is “OFF” and 1 is “ON”; equations used to calculate expression in each case are in the Supplementary Information. For the NAND, EQ and XOR gates we assume that when TFs act to reduce the effective rate of their target step, they do so by increasing the free energy of transition; as such, their effects on the rate are multiplicative (see Supplementary Information). The effect of repressor R on a rate is described here as $k_i(R) = \frac{k_i(0)}{(1+r_i)}$, where r_i is the ratio between the repressor concentration and the operator’s dissociation constant ($r_i \geq 0$).

Figure 3. Combinatorial regulation of the transcription cycle enables logical computations on out-of-sync nuclear localization events.

(A) Recent work has shown that transcription can be combinatorially controlled by transient bursts of transcription factor (TF) localization to the nucleus (Lin et al. 2015). Such localization events could be synchronous (upper panel), or asynchronous (lower panel), with a lag between them of time τ , as shown here for two TFs (orange and blue).

(B) Under the “recruitment” model of combinatorial transcriptional regulation, TFs must be simultaneously localized to the nucleus in order for a logical computation such as this AND gate to be possible. This is because cooperative physical interactions between the TFs in binding the DNA and/or recruiting RNA polymerase are required to stabilize the binding of RNAP to the promoter.

(C, D) In contrast, if transcription is combinatorially regulated by TFs acting on the transcription cycle, the TFs do not need to be on the DNA or even in the nucleus at the same time. We examine a transcription cycle with four effective rates (C, middle panel), which allows the implementation of a time encoded AND gate: An initial localization event of TF A (orange) to the nucleus accelerates the first rate, enabling the subsequent middle two steps in the cycle to occur; these two are slow but proceed without regulation from either TF. After a time delay (τ), TF B localizes to the nucleus and enables completion of the last step in

the cycle. With increasing time between the bursts of nuclear localization the fraction of cells responding to this temporally-encoded AND gate increases to 1.

(D) Temporal encoding of logical computations by regulating distinct steps in the transcription cycle provides a means for the promoter to distinguish the order in which inputs arrive. For this AND-gate example, a high level of expression cannot be achieved if (i) the time required for steps 1, 2 and 3 to occur is longer than the time that TFs A and B are simultaneously colocalized to the nucleus, or if (ii) B arrives before A.

Figure 4. Null models of combinatorial transcriptional control: independent action of activators

To interpret analog data from experiments that examine the effect of transcription factor (TF) number, type and strength on gene expression, we need a null model of combinatorial regulation. Here we define the null model as independent action of TFs. We simulate expression driven by two independently acting activators, using different assumptions about their functions at the promoter. To compare the outcomes, we plot expression relative to the null expectation in two specific limits. In the “recruitment” limit, the null expectation for expression (i.e. expression expected when TFs act independently) is that the *fold changes* in transcription due to each TF are *multiplicative*. On the x-axis we therefore plot deviation from this null, $\frac{Q_{AB}}{Q_A \times Q_B}$, where Q_{TF} is the fold change in expression over basal. In the “kinetic” limit of two activators working independently on a single slow step, by each working through different pathways, the expectation is that the *enhancement* of the rate by each TF is *additive*; on the y-axis we therefore plot deviation from this null expectation, $\frac{\Delta k_{TXN}(A,B)}{\Delta k_{TXN}(A) + \Delta k_{TXN}(B)}$.

(A) Under the “recruitment” hypothesis expression driven by two independently acting activators is expected to be the product of the fold changes due to each TF acting alone. All parameters in the accompanying equation were randomized, except for ω and ε , which represent interactions between the activators in binding DNA or contacting the promoter and were held at 1. K_A and K_B are the dissociation constants for TFs A and B; we randomized the parameters $\alpha = \frac{[A]}{K_A}$ and $\beta = \frac{[B]}{K_B}$ from 0.01-100. f_A and f_B represent the strength of interaction between each TF and the RNA polymerase and these values were randomized from 1-100.

(B) If either activator can catalyze the rate of the single slow step in transcription via different and independent pathways, the null expectation is that overall enhancement of the rate is the sum of the enhancement due to each TF acting alone. We show this numerically here, where the strength of each activator is randomized between 0-100 while the basal rate, k , is varied between 1-1000.

(C-E) If there is more than one slow rate in the transcription cycle, the null expectation for expression output is more complex and depends on the function of the activators, i.e. which step(s) in the transcription cycle each acts upon. Simulations show the range of possible gene expression outcomes from a simple cycle composed of two slow steps, k_1 and k_2 , controlled by two activators, A and B. The strength of enhancement by A and B of each rate is denoted by a_1, b_1, a_2 and b_2 , and these were varied between 0-100. This enhancement strength is determined here both by the intrinsic strength of the activator and by its concentration. k_1 and k_2 were randomized between 1-1000.

(C) Either activator can catalyze the rate k_1 , independently through different pathways. Their enhancement effects on k_1 sum together (see above), but because of the presence of the other rate, k_2 , their effects on the change in the overall rate of transcription are largely subadditive, tending to additivity in the limit in which $k_2 \gg k_1$.

(D) Two activators working exclusively on different steps in transcription.

(E) Both activators can work on both steps. For each individual step we assume, as in **(B and C)**, that they catalyze the rate independently through different pathways.

Box 2: Clarifying the meaning of synergy

The term ‘synergy’ arises frequently in discussions of how transcription factors (TFs) work together to affect gene expression. Synergy reflects deviations from a null expectation of independent action of components in a complex system (Salvador 2000). In the context of transcription, because the null expectation is not always rigorously defined, synergy is most often used to mean a “greater-than-additive” effect from multiple bound TFs. Although greater-than-additive effects may arise through various mechanisms described below, synergy itself should not be immediately inferred from it because it depends on the null expectation as we show. There are three main ways that greater-than-additive effects on expression might occur (Herschlag & Johnson 1993; Blau et al. 1996). Notably, two of these reflect null models of independent actions of TFs (and thus cannot actually reflect synergy). This is not an exhaustive list, nor are the mechanisms below mutually exclusive.

(i) **Cooperative DNA binding** due to direct or indirect recruitment interaction between TFs. In this case greater-than-additive effects should disappear at high concentrations of TFs (Carey et al. 1990; Lin et al. 1990).

(ii) **Simultaneous interaction of multiple TFs with the basal transcriptional machinery.** If TFs can simultaneously and independently stabilize the basal transcriptional machinery, their free energies will add, leading to a multiplicative effect on the rate of transcription (Veitia 2003; Herschlag & Johnson 1993). For most values of the individual activator contributions, this will be a greater-than-additive effect (it will not be when the fold-change in transcription for each activator is less than 2).

(iii) **Kinetic control of different steps of the transcription cycle.** If TFs stimulate more than one slow step in transcription then their combined effects will be greater than the sum of their individual actions. This is not the only kinetic means by which greater-than-additive expression could be achieved; for further discussion see Herschlag & Johnson (1993).

Science Application Box. Predictions arising from combinatorial control of the transcription cycle

(A) Distinguishing functional classes of transcription factors (TFs). Different roles of TFs in stimulating the transcription cycle may be distinguished by assessing the effect of a TF on a promoter that is functionally saturated—i.e. where the expression has plateaued. Adding a binding site for second TF that works on the same step in the cycle (orange) will not increase expression, while one that works on a different slow step will do so if that step is now limiting (red). Calculated expression levels are plotted for a two-state promoter with effective rates $k_1 = 1$, $k_2 = 10$ and activator strengths $a_1=10$, $b_1=2$, $c_2=2$.

(B) The role of the promoter in transcriptional control. Multimerizing binding sites for an activator (orange) that works on one step in a two-step cycle produces different expression depending on the relative rates of the two steps at different promoters (right, promoters A-D). Expression calculated for a two-state promoter with effective rates as shown in the plots to the right, and activator strength $a = 10$.

(C) Combinatorial control of transcription by two or more enhancers. Duplicated enhancers may drive little more expression than each enhancer does individually by virtue of responding to the same regulator(s) and thereby accelerating the same rates at the promoter to the same extent. If mutations arise in a duplicate that alter its function, the combined output of the two enhancers may be much higher. This is analogous to some sibling enhancer pairs, which are known to respond to different activators (Wunderlich et al. 2015; Staller et al. 2015). We speculate that these enhancers act differently on the transcription cycle

as a result; according to our model, they are thus likely to exhibit super-additive expression. Expression calculated for a transcription cycle with two effective rates, where $k_1 = k_2 = 1$ a.u. and the enhancement factors are $a_1 = a'_1 = 10$; $a_2 = a'_2 = 2$; $a'm_1 = 2$, $a'm_2 = 10$.

SUPPLEMENTARY INFORMATION

Calculating effective rates from a closed-loop cycle with reversible and irreversible steps.

Refers to Box 1 and first section of Results, and accompanies Figure S1.

In Box 1 we showed two examples of the ‘spectra’ of effective rates at a bacterial and a eukaryotic promoter. We calculated these effective rates, which incorporate the effects of reverse rates, using kinetic data from (Friedman & Gelles 2012) and (Stasevich et al. 2014)). Here we describe how we obtained these effective rates, which we refer to as ‘effective’ because they do not necessarily correspond to actual biochemical reactions.

The transcription cycle at the activator-dependent σ^{54} promoter of the *S. typhimurium* glnALG operon is comprised of reversible and irreversible transitions between four principal states (Friedman & Gelles 2012) (Figure S1A). State 1 is the unbound promoter, states 2 and 3 are closed complexes, and state 4 is the open complex; mRNA is produced from the elongation step (rate k_{41}). The forward and reverse rates connecting the four states were measured using multiwavelength single-molecule fluorescence colocalization (CoSMoS) (Friedman & Gelles 2012). The rate of RNAP association with the promoter, k_{12} , was determined to be $2.1 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$; we assumed that RNAP exists at $\sim 1 \mu\text{M}$ in the cell, giving a rate of rate of 21 s^{-1} , which we use in these calculations. In the absence of the bacterial activator, k_{34} is $< 6 \times 10^{-6}$; this rate increases 300-fold to $1.93 \times 10^{-3} \text{ s}^{-1}$ when the activator NtrC is added along with ATP, but the rates spectrum shown here is in the absence of the activator.

This four-step cycle (Figure S1B) is comprised of ten constituent ‘pseudo’-times that are summed together to calculate the total time taken to complete the cycle (T_{TXN} ; Equation 3, main text). These terms do not correspond to real biochemical reactions. For example, the effective time the promoter spends trying to move past the first state in the cycle, T_1^* , is the sum of four ‘pseudo’ times τ_1 , $\tau_1 R_2$, $\tau_1 R_2 R_3$ and $\tau_1 R_2 R_3 R_4$, where R is the rejection rate (Equation 1, main text), which is equivalent to saying that there are four irreversible reactions in series. The ten pseudo-steps are diagrammed in Figure S1B and their rates (the inverse of the time taken to complete each pseudo-step) are plotted in Figure S1C, color-coded by the state transition to which they belong.

For the eukaryotic transcription cycle example we used kinetic measurements of rates in the transcription cycle made by Kimura and colleagues on a glucocorticoid receptor activated tandem gene array in a mouse cell line. These authors developed a means of tracking specific histone modifications and RNAP phosphorylation in living cells and used these to measure the association of RNAP with the promoter in the presence of GR (state 1 \rightarrow state 2, Figure S1D), promoter escape (state 2 \rightarrow state 3; associated with phosphorylation of serine 5 of the RNAP C-terminal domain), and escape into sustained elongation (state 3 \rightarrow state 1; associated with ser2 phosphorylation of the C-terminal domain). The corresponding forward and reverse rates are detailed in Figure S1D.

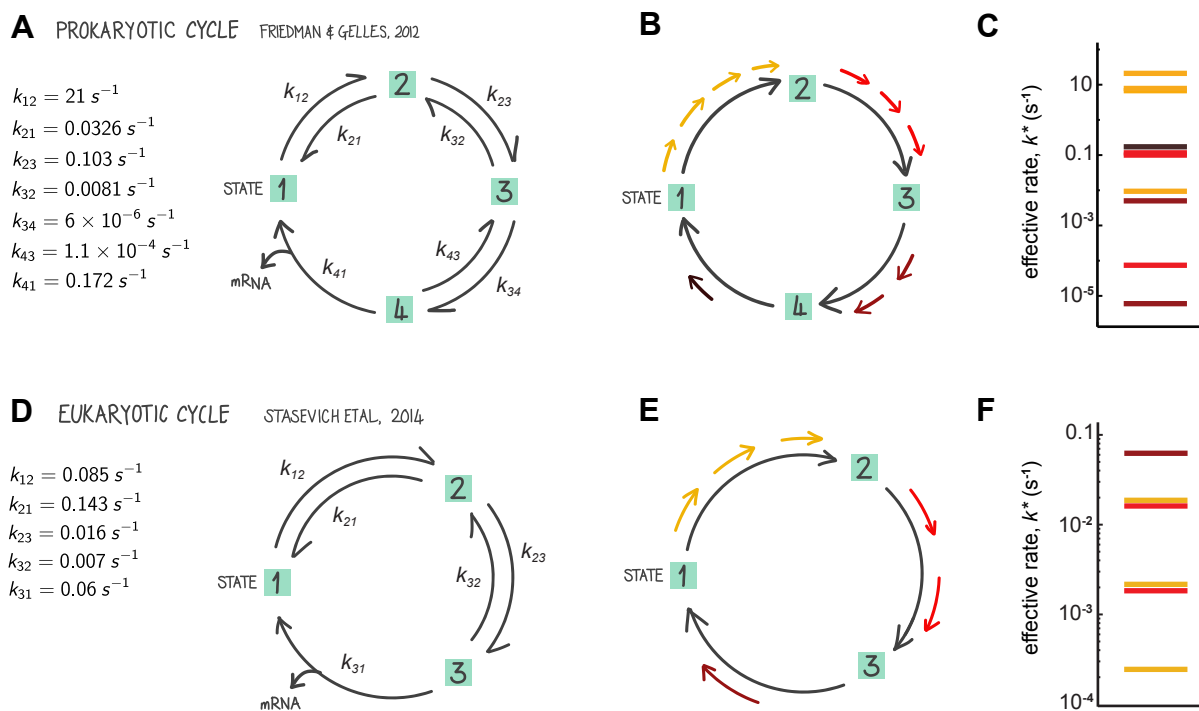


Figure S1. Calculating effective rates from a closed-loop cycle with reversible and irreversible steps.

This figure accompanies the explanation in the text of the Supplementary Information and that in the Results section. The forward and reverse rates for a prokaryotic and a eukaryotic transcription cycle were obtained from Friedman & Gelles (2012) and Stasevich et al. (2014), respectively. These are listed next to a simplified representation of each cycle in (A) and (D). As described in the accompanying text, we calculated the effective forward rates for each cycle; these take into account the effect of reverse rates. For example, in (A-C) the promoter transitions between four states; the total time taken to complete the cycle, T_{TOTAL} , is the sum of the times taken to complete each step. For the four-step cycle in question, T_{TOTAL} comprises ten terms (see Results), and the inverse of each term corresponds to an 'effective' rate. We refer to them as 'effective' because they don't necessarily correspond to a real biochemical reaction since they incorporate the effects of all the reverse rates in the cycle). In (C) and (F) we plot the effective rates (k^*) for each cycle.

Assumptions made in examples of logic gate formation through control of the transcription cycle

Table S1: Parameter values used in Figure 2

	AND	OR	NOR	NAND	EQ	XOR
k_1	1	1	100	100	100	1
k_2	1	100	100	100 000	100	10 000
a_1	100	1000	0.01	0	100	100
b_1	0	1000	0	0	0.01	100
a_2	0	0	0	0.01	0.01	0.01
b_2	100	0	0.01	0.01	100	0.01

Supplementary Table 2: Calculating rates of transcription in Figure 2

	AND	OR	NOR
$k_{TXN}(0)$	$\frac{k_1 \times k_2}{k_1 + k_2}$	$\frac{k_1 \times k_2}{k_1 + k_2}$	$\frac{k_1 \times k_2}{k_1 + k_2}$
$k_{TXN}(A)$	$\frac{(a_1 + 1)k_1 \times k_2}{(a_1 + 1)k_1 + k_2}$	$\frac{(a_1 + 1)k_1 \times k_2}{(a_1 + 1)k_1 + k_2}$	$\frac{(a_1)k_1 \times k_2}{(a_1)k_1 + k_2}$
$k_{TXN}(B)$	$\frac{k_1 \times (b_2 + 1)k_2}{k_1 + (b_2 + 1)k_2}$	$\frac{(b_1 + 1)k_1 \times k_2}{(b_1 + 1)k_1 + k_2}$	$\frac{k_1 \times (b_2)k_2}{k_1 + (b_2)k_2}$
$k_{TXN}(A,B)$	$\frac{(a_1 + 1)k_1 \times (b_2 + 1)k_2}{(a_1 + 1)k_1 + (b_2 + 1)k_2}$	$\frac{(a_1 + b_1 + 1)k_1 \times k_2}{(a_1 + b_1 + 1)k_1 + k_2}$	$\frac{(a_1)k_1 \times (b_2)k_2}{(a_1)k_1 + (b_2)k_2}$

	NAND	EQ	XOR
$k_{TXN}(0)$	$\frac{k_1 \times k_2}{k_1 + k_2}$	$\frac{k_1 \times k_2}{k_1 + k_2}$	$\frac{k_1 \times k_2}{k_1 + k_2}$
$k_{TXN}(A)$	$\frac{k_1 \times (a_2)k_2}{k_1 + (a_2)k_2}$	$\frac{(a_1 + 1)k_1 \times (a_2)k_2}{(a_1 + 1)k_1 + (a_2)k_2}$	$\frac{(a_1 + 1)k_1 \times (a_2)k_2}{(a_1 + 1)k_1 + (a_2)k_2}$
$k_{TXN}(B)$	$\frac{k_1 \times (b_2)k_2}{k_1 + (b_2)k_2}$	$\frac{(b_1)k_1 \times (b_2 + 1)k_2}{(b_1)k_1 + (b_2 + 1)k_2}$	$\frac{(b_1 + 1)k_1 \times (b_2)k_2}{(b_1 + 1)k_1 + (b_2)k_2}$
$k_{TXN}(A,B)$	$\frac{k_1 \times (a_2 \times b_2)k_2}{k_1 + (a_2 \times b_2)k_2}$	$\frac{b_1(a_1 + 1)k_1 \times a_2(b_2 + 1)k_2}{b_1(a_1 + 1)k_1 + a_2(b_2 + 1)k_2}$	$\frac{(a_1 \times b_1 + 1)k_1 \times (a_2 \times b_2)k_2}{(a_1 \times b_1 + 1)k_1 + (a_2 \times b_2)k_2}$

There are many potential ways in which the logic gates in Figure 3 might be achieved mechanistically. We focus only on examples of necessary basal rates in the transcription cycle and changes from them. In the scenarios in Figure 3 we assumed the following. AND gate: Each TF independently catalyzes a different rate. OR gate: Either TF can independently catalyze the slower rate, therefore their effects on this rate add together. NOR gate: Each TF acts as a repressor on a different rate. The effect of repressor R on a rate is described here as $k_i(R) = \frac{k_i(0)}{(1+r_i)}$, where r_i is the ratio between the repressor concentration and the operator's dissociation constant ($r_i \geq 0$). NAND gate: Both TFs can independently repress the faster of the two rates. We assume for the NAND and EQ gates that repressors act by increasing the free energy required to complete a step. Therefore, when they work independently on the same step, their combined effects on this rate of this step are multiplicative. EQ gate: We assume that each TF activates one rate and represses the other. This might for example occur if, while carrying out its activating function at rate 1, the TF occludes recruitment of a key enzyme necessary for step 2 to proceed. XOR gate: Both TFs increase the slower rate while decreasing the faster rate. When both are present, their combined effects on each rate are multiplicative. This could occur if the TFs act by stabilizing or destabilizing a given protein complex.

Under the recruitment model, expression is multiplicative when TFs act independently

The null expectation for expression controlled by two activators acting independently under the recruitment model is multiplicativity (Bintu et al. 2005b; Veitia 2003). This has been derived analytically, which we recap below to accompany Figure 4A, but we also demonstrate it numerically in that Figure.

In the recruitment model, activators recruit RNAP to the promoter by lowering the free energy of the bound RNAP state. The rate limiting step is the recruitment of RNAP because k_{OFF} is extremely fast, such that RNAP falls off the promoter rapidly before getting the chance to fire. Activators work by reducing this off-rate. We assume that activators A and B have affinities described by the dissociation constants $K_A = e^{-\Delta G_A}$ and $K_B = e^{-\Delta G_B}$, where ΔG_A and ΔG_B represent the binding free energies for the two activators. We consider activators A and B to be present at concentrations $[A]$ and $[B]$. When bound at the same time, both activators stabilize each other by reducing their dissociation constants by a factor $\omega = e^{-\Delta G_{AB}}$. A and B also lower the free energy of the bound RNAP state by ΔG_{AP} and ΔG_{BP} respectively, and they thus decrease the dissociation constant of RNAP by magnitudes $f_A = e^{-\Delta G_{AP}}$ and $f_B = e^{-\Delta G_{BP}}$. If the two TFs contact the RNAP independently from each other, the net reduction in free energy of the bound RNAP will be the sum of the reductions caused by each TF. If they do not act independently, the extent of the deviation is described by a parameter $\varepsilon = e^{-g_{AB}}$, so that $f_{AB} = e^{-\Delta G_{ABP}} = e^{-(\Delta G_{AP} + \Delta G_{BP} + g_{AB})} = f_A \times f_B \times \varepsilon$. When only the binding site of A is present (and defining for convenience the unitless parameter $\alpha = \frac{[A]}{K_A}$), the fold-change activation takes the form $Q_A = \frac{(1 + \alpha f_A)}{(1 + \alpha)}$. Likewise, and defining $\beta = \frac{[B]}{K_B}$, the fold-change activation when only the binding site of B is present is $Q_B = \frac{(1 + \beta f_B)}{(1 + \beta)}$. The fold-change activation of when both binding sites are present takes the form $Q_{AB} = \frac{(1 + \alpha f_A + \beta f_B + \alpha \beta f_A f_B \omega \varepsilon)}{(1 + \alpha + \beta + \alpha \beta \omega)}$.

These results can be easily derived from the model, but they can be also found elsewhere (Bintu, Buchler, Garcia, Gerland, Hwa, Kondev, Kuhlman, et al. 2005; Bintu, Buchler, Garcia, Gerland, Hwa, Kondev & Phillips 2005). In Figure 4A, since we are interested in the null situation of independent action of

activators (i.e. $g_{AB} = 0$, $\Delta G_{AB} = 0$), we set $\varepsilon = 1$, $\omega = 1$ in the model, and randomized all of the other parameters in the above equation for Q_A , Q_B , and Q_{AB} .

Under the recruitment model, additivity is rare

Under the recruitment model, in order to get additive effects between two binding sites when the activators act independently to recruit RNAP, it is necessary that all of the following conditions are fulfilled (refer to the equation in the previous section): (i) $\alpha \ll 1$, (ii) $\beta \ll 1$, (iii) $f_A \times \alpha \gg 1$, (iv) $f_B \times \beta \gg 1$ and (v) either $\omega = 0$ or $\varepsilon = 0$; i.e. the activators bind to their binding sites only rarely but have very strong effects on the RNAP when they are bound, and they cannot simultaneously contact RNAP. Therefore, while additivity is in principle possible under the recruitment mechanism, it can only occur under very restricted conditions; therefore the recruitment mechanism is unlikely to yield additive behavior.

SUPPLEMENTARY FIGURE AND TABLE LEGENDS

Table S1. Parameter values used to generate logic gates in Figure 2.

Table S2. Calculating rates of transcription for logic gates in Figure 2.

Equations used to calculate the rate of transcription for the example logic gates shown in Figure 2. We consider a simple cycle comprised of two steps, described by rate constants, k_1 and k_2 . $k_{TXN}(0)$, overall rate of transcription in the absence of activators; $k_{TXN}(A)$, transcription rate in the presence of activator A; $k_{TXN}(AB)$, transcription rate when both activators A and B are present. a_1 and a_2 represent the enhancement of rates k_1 and k_2 , respectively, by activator A. When an activator has no effect on a rate we have left it out of the equations for clarity.

Figure S1. Calculating effective rates from a closed-loop cycle with reversible and irreversible steps.

This Figure accompanies the rates “spectra” in Box 1 and the explanation of calculating the transcription rate from a cycle with no internal loops in the Results section. The forward and reverse rates for a prokaryotic and a eukaryotic transcription cycle were obtained from Friedman & Gelles (2012) and Stasevich et al (2014), respectively. These are listed next to a simplified representation of each cycle in (A) and (D). As described in the accompanying text, we calculated the effective forward rates for each cycle; these take into account the effect of reverse rates. For example, in (A-C) the promoter transitions between four states; the total time taken to complete the cycle, T_{TOTAL} , is the sum of the times taken to complete each step. For the four-step cycle in question, T_{TOTAL} comprises ten terms (see Results), and the inverse of each term corresponds to an ‘effective’ rate. We refer to them as ‘effective’ because they don’t necessarily correspond to a real biochemical reaction since they incorporate the effects of all the reverse rates in the cycle. In (C) and (F) we plot the effective rates (k^*) for each cycle.