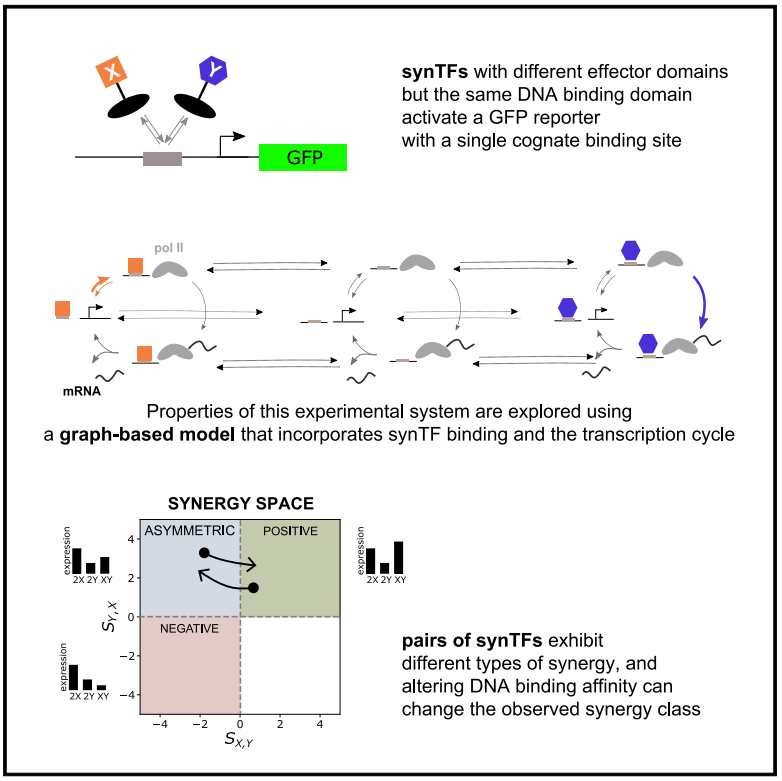
Cell Systems

Report

Transcriptional kinetic synergy: A complex landscape revealed by integrating modeling and synthetic biology

Graphical abstract



Authors

Rosa Martinez-Corral, Minhee Park, Kelly M. Biette, ..., Ahmad S. Khalil, Jeremy Gunawardena, Angela H. DePace

Correspondence

angela_depace@hms.harvard.edu

In brief

Martinez-Corral, Park, Biette et al. investigate kinetic synergy in which transcription factors enhance each other's effect through functional complementarity. To eliminate binding cooperativity, the authors focus on a synthetic scenario where transcription factors bind to a shared site. The combination of modeling and experiments reveals a complex landscape in which binding and function determine synergy.

Highlights

- TFs with complementary activities (kinetic roles) can synergize (kinetic synergy)
- A synthetic system where TFs bind to the same site reveals kinetic synergy
- Synergy depends on the binding kinetics and kinetic roles of the TFs

Martinez-Corral et al., 2023, Cell Systems *14*, 324–339
April 19, 2023
https://doi.org/10.1016/j.cels.2023.02.003

CellPress

CellPress
OPEN ACCESS

Cell Systems

Report

Transcriptional kinetic synergy: A complex landscape revealed by integrating modeling and synthetic biology

Rosa Martinez-Corral,[1,5] Minhee Park,[2,3,6,5] Kelly M. Biette,[1,7,5] Dhana Friedrich,[1,8] Clarissa Scholes,[1,9] Ahmad S. Khalil,[2,3,4] Jeremy Gunawardena,[1] and Angela H. DePace[1,10,*]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[2]Biological Design Center, Boston University, Boston, MA 02215, USA
[3]Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA
[4]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
[5]These authors contributed equally
[6]Present address: Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea
[7]Present address: Mercy BioAnalytics, Natick, MA 01760, USA
[8]Present address: Bayer AG Pharmaceuticals, 51368 Leverkusen, Germany
[9]Present address: Mammoth Biosciences, Brisbane, CA 94005, USA
[10]Lead contact
*Correspondence: angela_depace@hms.harvard.edu
https://doi.org/10.1016/j.cels.2023.02.003

SUMMARY

Transcription factors (TFs) control gene expression, often acting synergistically. Classical thermodynamic models offer a biophysical explanation for synergy based on binding cooperativity and regulated recruitment of RNA polymerase. Because transcription requires polymerase to transition through multiple states, recent work suggests that "kinetic synergy" can arise through TFs acting on distinct steps of the transcription cycle. These types of synergy are not mutually exclusive and are difficult to disentangle conceptually and experimentally. Here, we model and build a synthetic circuit in which TFs bind to a single shared site on DNA, such that TFs cannot synergize by simultaneous binding. We model mRNA production as a function of both TF binding and regulation of the transcription cycle, revealing a complex landscape dependent on TF concentration, DNA binding affinity, and regulatory activity. We use synthetic TFs to confirm that the transcription cycle must be integrated with recruitment for a quantitative understanding of gene regulation.

INTRODUCTION

The regulation of transcription is a finely controlled process central to biology, biomedicine, and bioengineering applications. At its core are transcription factors (TFs), proteins that bind specific sites on the DNA and directly or indirectly modulate the binding and activity of the RNA polymerase complex. In eukaryotes, multiple TFs of the same and distinct types collaborate to drive transcription through binding to gene regulatory regions called enhancers and promoters.[1] Such "combinatorial control" enables binding and response specificity[2,3] and expands the regulatory capabilities of the finite set of TFs encoded by an organism. A wealth of studies have characterized TF binding sites and binding profiles in model genes, genomes, and random sequences (e.g., Smith et al.,[4] Vandel et al.,[5] and Inukai et al.[6]). In turn, a long-standing goal of biomedicine and synthetic biology has been to exploit this type of information to anticipate the effect of mutations on regulation, to develop new and more refined pharmacological interventions, and to design next-generation synthetic circuits with more precise and robust functions. However, this is still a difficult task, in part because of the non-independent effects of the TFs that control a given gene.[7–11]

When TFs interact to regulate transcription, the response to a combination of TFs is often not simply predicted by the responses to each of the TFs alone. Some models indicate that in the absence of interactions between TFs or sites, their combined effect should be the addition of the individual outputs, and synergy has been used to refer to deviation from this additive expectation.[12–14] Under other models, synergy is manifested as multiplicativity in the response.[15] Alternatively, the term "synergy" has been used to refer to nonlinear response to increasing TF concentrations,[16] binding cooperativity (below), or a special form of it.[17,18] Here, we use this term to refer to an increase in the expression output under two TFs in comparison with their individual effects, quantified by a functional measure tailored to our experimental system (results).

Synergy has commonly been understood through the lens of recruitment models of transcription in which the role of TFs is to regulate the binding of the RNA polymerase to the gene.[19] Thermodynamic models of gene regulation offer a biophysical

grounding for this view.[15,20,21] These models assume that TFs and polymerase bind to the DNA under thermodynamic equilibrium conditions. The free energy of each state determines its steady-state probability according to the Boltzmann distribution, and the transcription rate is treated as a function of the states of binding of the system. Synergy then emerges from direct or indirect cooperative binding interactions in which TFs enhance or reduce each other's binding and that of the RNA polymerase to the DNA (e.g., Vashee et al.,[22] Ambrosetti et al.,[23] Spitz et al.,[24] Frank et al.,[25] Goldstein et al.,[26] and Estrada et al.[27]). Mechanistically, this can result from direct protein-protein interactions between adjacently bound molecules,[28] indirect interactions through a shared molecule or complex such as Mediator, [12,29,30] or allosteric mechanisms[31] mediated by nucleosomes[32] or DNA.[33]

Beyond the recruitment of RNA polymerase to the gene, it is well known that eukaryotic transcription is a multi-step process that is regulated at many points. Accordingly, it has been suggested that transcriptional regulation should be understood in terms of a transcription cycle,[34] involving the displacement of nucleosomes at the start site, post-translational modification of histones,[35–37] assembly of the transcriptional machinery, and post-translational modifications that regulate its activity and elongation rate.[38,39] In agreement with this view, RNA polymerase has been found to be already bound to many inactive genes, suggesting that, under certain scenarios, activation does not rely on regulating polymerase recruitment but rather on modulating a subsequent step.[40] Besides shifting the focus away from the recruitment of the RNA polymerase, this view also implies non-equilibrium behavior, given that ATP-dependent nucleosome remodeling and post-translational modifications involve energy dissipation. In this case, the steady-state behavior of the system is determined by the individual rates of the various transitions. This is in contrast to the equilibrium situation of thermodynamic models in which only the ratios between the forward and backward rates matter for determining the steady state of the system.[41]

Under this kinetic view, the possibility of "kinetic synergy" was theoretically proposed. Imagine the simplest case in which transcription is regulated by two steps and two TFs have different biochemical functions (or, more generally, "kinetic roles," because the detailed biochemical reactions regulated by a TF are generally not known[42]), such that one TF preferentially enhances one step and the other TF preferentially enhances the other. Then, when the two TFs are present together, they can enhance each other's effects and thus generate synergy.[13,14] Importantly, TFs with complementary kinetic roles would enable synergy to emerge even in the absence of cooperative binding between TFs on the DNA; the TFs would not even need to be simultaneously present at the regulatory site.

Multiple lines of evidence make kinetic synergy very plausible. First, experimental work has shown that transcriptional activators can increase gene expression by different mechanisms. TF activation domains were found to either stimulate transcription initiation, elongation, or both.[43] More recent studies have continued to reveal that TFs use diverse mechanisms to regulate transcription and affect distinct steps of the transcription cycle (e.g., Harden et al.,[42] Fu et al.,[44] Rahl et al.,[45] Baluapuri et al.,[46] and Bell et al.[47]). Along the same lines, differences in RNA polymerase II pausing after treatment with E2 or TNF-alpha signals have been attributed to the TFs downstream (ERα and NF-κB) acting on different transitions that regulate their target genes.[48] Moreover, comparisons between regulation driven by homogeneous or heterogeneous sets of TFs have shown that heterogeneous sets often drive higher expression levels.[4,49,50] In line with this, evidence of synergistic activation between the viral activator VP16 and selected chromatin regulators in a reporter system has been observed.[51] Similarly, the activity of many *Drosophila* TFs and cofactors has been found to be highly context-dependent,[52] suggesting that activation may require a particular combination of biochemical mechanisms.

Despite these observations, it is experimentally challenging to assess kinetic synergy given the difficulty of disentangling it from cooperative DNA-binding interactions between TFs. On the theoretical side, there have been few tools to reason about kinetic synergy on biophysical grounds. As a first step, a recent theoretical study by our group showed that, in a similar way to binding cooperativity, kinetic synergy can implement logical and analog computations,[14] and that it can generate a wide diversity of input/output relationships. However, in a similar way to other modeling work that considers transcription as a multi-step process (e.g., Hansen et al.,[36] Suter et al.,[53] and Rybakova et al.[54]), that model did not explicitly account for TF binding and instead represented it indirectly through the effect of the TFs on the transition rates of the system. To our knowledge, there have been few attempts to explicitly model the interplay between TF binding, polymerase recruitment, and progression over the transcription cycle. A model that explicitly incorporated binding and transitions over the cycle has been proposed,[55] but it assumed a timescale separation between TF binding and the rest of the processes, with quasi-equilibrium in TF binding. However, both TF residence times and the half-life of certain biochemical steps in the transcription cycle may occur on similar timescales, on the order of several seconds or a few minutes (STAR Methods, biologically plausible ranges for parameters), calling for more general models that bring together the binding-centered view of recruitment models with the regulation of the transcription cycle.

Here, we exploit the graph-based linear framework (below) to propose a model of transcriptional control that explicitly accounts for TF binding and the regulation of polymerase recruitment, as well as the progression over the transcription cycle. In order to disentangle kinetic synergy from binding cooperativity, we focus on the emergence of synergy between TFs binding to a single, shared site. This scenario eliminates the possibility of TFs simultaneously bound to the DNA, thus removing cooperative binding between TFs. Experimentally, we build this system using engineered TFs in which the activation domains of a set of functionally diverse mammalian TFs are fused to a computationally designed zinc-finger (ZF) DNA-binding domain predicted to bind only to an artificial site upstream of a reporter (Figure 1A).[51,56–58] We propose a measure of synergy in which we compare the expression output when both TFs are present with that when only one of them is present. By exploring the synergistic behavior of the model in parameter space, we find that a diversity of behaviors can emerge in this scenario, for which we find experimental evidence. Our model reveals a complex synergy landscape, shaped by the interplay between the kinetic role of the TFs and their binding kinetics. This highlights the

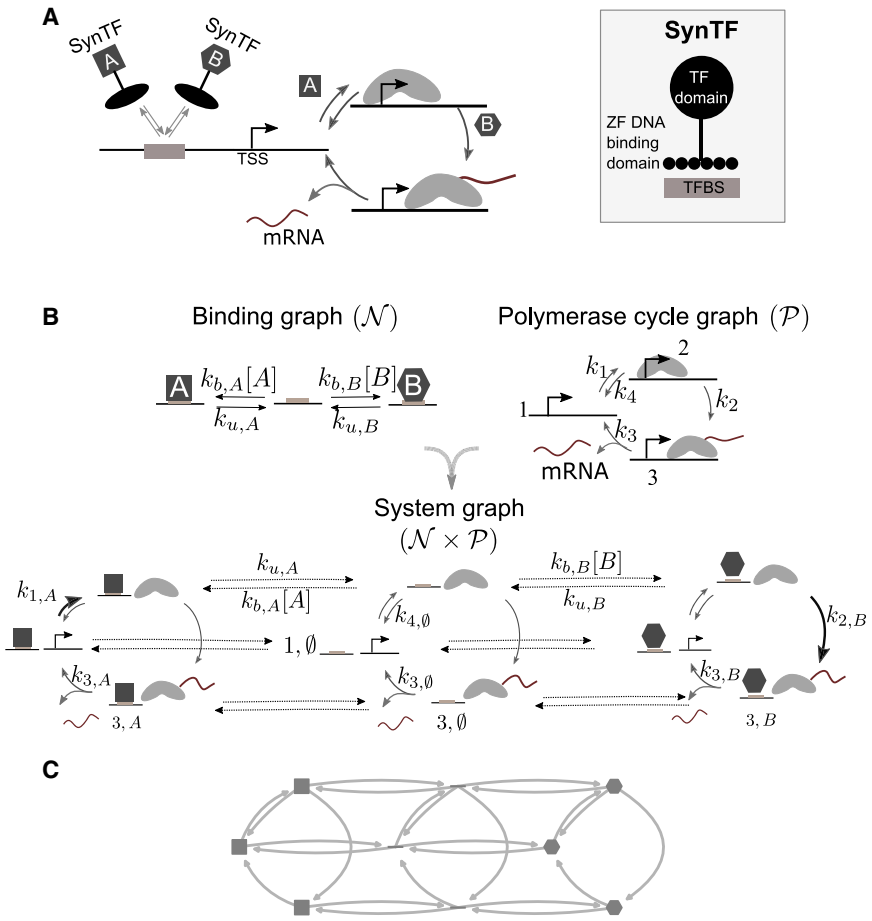


Figure 1. A model for kinetic synergy between two TFs sharing a single DNA binding site

(A) Cartoon schematizing the strategy of this work to examine kinetic synergy: two synthetic TFs (synTFs) regulate a reporter (not shown) through a shared binding site (TFBS). As an example, TF *A* controls the first step in the transcription cycle, and TF *B* controls the second step.

(B) Model used in this work. The structure (vertices and edges) of the graph of the system ($\mathcal{N} \times \mathcal{P}$) corresponds to the structure that results from the graph product between the binding graph $\mathcal{N}$ and the 3-state polymerase cycle graph $\mathcal{P}$. Only a subset of nodes and edges are labeled, for clarity. The horizontal edges from the central cycle to the outer cycles denote binding of each of the TFs, and the reverse edges denote unbinding. The three cycles allow us to account for the effect of the TFs, because the rates can be different depending upon the state of the binding site. As an example, the darker arrows denote the activator effect of *A* and *B* on the first and second transitions, respectively ($k_{1,A} > k_{1,\emptyset}$, $k_{2,B} > k_{2,\emptyset}$).

(C) Schema of the full graph, as used in Figure 2.

relevance of considering genomic context, binding, and biochemical function together when characterizing TFs and illuminates how functional interactions between TFs may contribute to eukaryotic transcriptional control.

RESULTS

Mathematical model

We study how kinetic synergy emerges in a scenario where two TFs bind to a shared site in a regulatory sequence, such that only one TF can be specifically bound at any given time. Figure 1A schematizes this situation for a general 3-state transcription cycle in which TF *A* promotes the first step (illustrated as the assembly of the RNA polymerase complex) and TF *B* promotes a process downstream.

In order to model this system, we exploit the linear framework formalism, a graph-based approach to Markov processes that can be used to model a diversity of biological processes in a biophysically realistic and mathematically tractable way.[59–61] We have previously applied this framework to study how binding interactions between TFs modulate gene expression by implicitly averaging over the states of the polymerase cycle.[27,62,63] In contrast, in a previous study of kinetic synergy, we modeled the effect of TFs on a detailed transcription cycle but effectively combined their binding with their enzymatic effects.[14] Here, we propose a model that unifies both approaches and does not make assumptions about the binding reactions being on a different timescale than the polymerase cycle reactions, improving previous approaches in the literature[55] (STAR Methods, modelling details and the linear framework).

The system is represented by a linear framework graph, whose construction is described in Figure 1B. The vertices of such graphs are the biological states of interest, and the edges are the transitions between them, assumed to follow Markovian dynamics. The edge labels are the infinitesimal transition rates, with dimensions of (time)$^{-1}$. The graph $\mathcal{N} \times \mathcal{P}$ in Figure 1B is the graph product of a binding graph, $\mathcal{N}$, and a polymerase cycle graph, $\mathcal{P}$. The binding graph, $\mathcal{N}$, consists of a binding site that can either be empty, bound by TF *A*, or bound by TF *B*. For the polymerase cycle graph, $\mathcal{P}$, we consider the simplest cycle, with 3 vertices (indexed 1, 2, and 3). These vertices could be interpreted in terms of the empty transcription start site (TSS), assembled RNA polymerase, and C-terminal phosphorylated or elongating polymerase, although mapping onto specific states of the cycle is not required to interpret the results. The first transition is assumed to be reversible and the other two irreversible in agreement with the macroscopic irreversibility of mRNA synthesis. mRNA is assumed to be produced when the system transitions from state 3 to state 1. The graph product, $G_A \times G_B$, of any two graphs, G_A and G_B, is constructed as follows—the vertices of $G_A \times G_B$ are ordered pairs, (i,j) of vertices, where i is a vertex in G_A and j is a vertex in G_B. If $i \to k$ is an edge in G_A, then there is an edge $(i,j) \to (k,j)$ in $G_A \times G_B$ for all vertices j in G_B. Similarly, if $j \to h$ is an edge in G_B, there is an edge $(i,j) \to (i,h)$ in $G_A \times G_B$ for all vertices i in G_A. These are all the edges in $G_A \times G_B$. Figure 1B depicts the graph product $\mathcal{N} \times \mathcal{P}$, but it is important to

note that the product construction only defines the structure of $\mathcal{N} \times \mathcal{P}$, as given by its vertices and edges, and does not specify its edge labels. This allows us to introduce assumptions about the interplay between the TFs and the polymerase cycle by choosing appropriate edge labels in $\mathcal{N} \times \mathcal{P}$, as explained further below.

TF binding on-rates ($k_{b,X}, X \in \{A,B\}$, horizontal from the central cycle to the right and left) have dimensions of (concentration × time)$^{-1}$, and binding off rates ($k_{u,X}, X \in \{A,B\}$) have dimensions of (time)$^{-1}$. The genomic context is modeled by the values of the basal rates over the polymerase cycle in the absence of TFs (central cycle). To incorporate the effect of a TF on a given transition, we assume that the TF only has effect while it is bound. The effect is then incorporated into the edge label for that transition, making it different for the cycle in which the TF is bound than for the basal cycle. As an example, the darker arrows on the left and right cycles in Figure 1B, $\mathcal{N} \times \mathcal{P}$, represent the activating effect of A and B on the first and second transitions, respectively. In this case, $k_{1,A} > k_{1,\varnothing}$, and $k_{2,B} > k_{2,\varnothing}$. Similarly, repression could be included as well by a smaller value for a transition rate than the corresponding basal rate. For simplicity, here we examine synergy between "pure" activators only, defined by not decreasing the clockwise rates ($k_{1,X} \geq k_{1,\varnothing}$, $k_{2,X} \geq k_{2,\varnothing}$, $k_{3,X} \geq k_{3,\varnothing}$, $X \in \{A,B\}$) and not increasing the counterclockwise rate ($k_{4,A} \leq k_{4,\varnothing}$, $k_{4,B} \leq k_{4,\varnothing}$).

We interpret the system in probabilistic terms and assume each vertex of the graph holds the probability of the system being in that state. The transition rates then determine the time-evolution of the probabilities according to the master equation, which eventually reach a steady state (STAR Methods, modelling details and the linear framework). Moreover, we assume first-order mRNA degradation. By taking the mRNA degradation rate as a constant that normalizes the transition rates, the steady-state mRNA at a given concentration of A and B ($m(A,B)^*$) is given by:

$$m(A,B)^* = k_{3,\varnothing} P^*_{3,\varnothing}(A,B) + k_{3,A} P^*_{3,A}(A,B) + k_{3,B} P^*_{3,B}(A,B) \quad \text{Equation (1)}$$

where $P^*_{3,\varnothing}(A,B)$, $P^*_{3,A}(A,B)$, $P^*_{3,B}(A,B)$ are the steady-state probabilities of state $3_\varnothing$, 3_A, 3_B at concentrations A and B of the TFs, and the rates are normalized by the mRNA degradation rate (STAR Methods, modelling details and the linear framework). Given that we only consider the steady-state behavior of the system, we use the same symbols to refer to the original rates and the normalized rates in order to avoid excessive notation. In the remainder of the paper, the rates will always be normalized.

The focus of our analysis is to compare this quantity when both TFs are present to that when only one is present and the other is at concentration 0 (synergy, below). Note that when only one or none of the TFs is present, m^* can be computed in the same way. In that case, the steady-state probabilities for those states corresponding to the absent TF being bound will be 0, and the rest will be redistributed according to the parameter values. The value of m^* in the absence of TFs ($m^*(0,0)$) corresponds to basal expression. For simplicity, the absence of a TF from the mathematical expressions below means it is at concentration 0.

A measure of synergy

Our interest is to understand how synergy emerges in this system. If two TFs act on more than one step in the cycle, the overall effect may not be greater than additive even if they interact kinetically.[14] This exemplifies that considering addition as a null expectation against which to define synergy, as has often been done in the literature, is model specific. Here, we propose a synergy measure that is particularly suited to our single binding site scenario and corresponding synthetic experimental approach in which TFs share the same DNA binding domain (below). We assume that the DNA binding domain determines the binding kinetics of the TFs to the DNA, so that both TFs bind and unbind to the DNA with the same kinetics and only differ in their kinetic roles. In this case, overall TF binding to the DNA binding site should remain unchanged when either one or two TFs are present at a fixed total TF concentration. Therefore, we can quantify synergy by comparing the steady-state expression when both TFs are present ($m^*(A,B)$) with the steady-state expression when either of them is alone but at twice as much concentration ($m^*(2A)$, $m^*(2B)$) (same total TF). Enhanced expression in combination with respect to the strongest TF (the TF with a higher level of expression on its own), or reduced with respect to the weakest, must arise as a result of functional interactions among the TFs over the cycle and corresponds to kinetic synergy.

Positive synergy corresponds to higher expression in combination as compared with individually and can be regarded as "canonical" synergy in the sense of enhanced expression in combination: expression is greater than that of the strongest TF even if half the molecules are substituted by those of a weaker TF. We note, however, that the output does not have to be greater than additive to be considered positive synergy. Negative synergy corresponds to lower expression in combination, with expression lower than that of the weakest TF alone. Asymmetric synergy results when expression is increased only with respect to the weakest TF. In this case, it may be unclear whether there are any synergistic interactions. Potentially, these can still be detected depending on the extent to which the expression is reduced or increased with respect to the strongest or weakest TF, respectively. Thus, we propose to quantify synergy as a point in 2D, by comparing the effects of adding one TF to the other. This is quantified by $S_{A,B}$ (effect of B on A) and $S_{B,A}$ (the effect of A on B) as follows:

$$S_{A,B} = \log_2\left(\frac{m^*(A,B)}{m^*(2A)}\right) \quad \text{(Equation 2)}$$

$$S_{B,A} = \log_2\left(\frac{m^*(A,B)}{m^*(2B)}\right) \quad \text{(Equation 3)}$$

If A is taken to be the strongest TF, positive (green), asymmetric (blue), and negative (red) synergy map to 3 quadrants of a two-dimensional synergy space, as depicted in Figure 2A-left.

In principle, we could also define synergy at various other ratios of TFs, keeping the total concentration fixed, and we could do this for various total TF concentrations. This would define, for each pair of TFs with their own kinetic roles and shared DNA binding kinetics, a two-dimensional map (Figure 2A-right), where each axis corresponds to a given total TF concentration. The black dot in Figure 2A-right denotes the 0.5:0.5 ratio of our

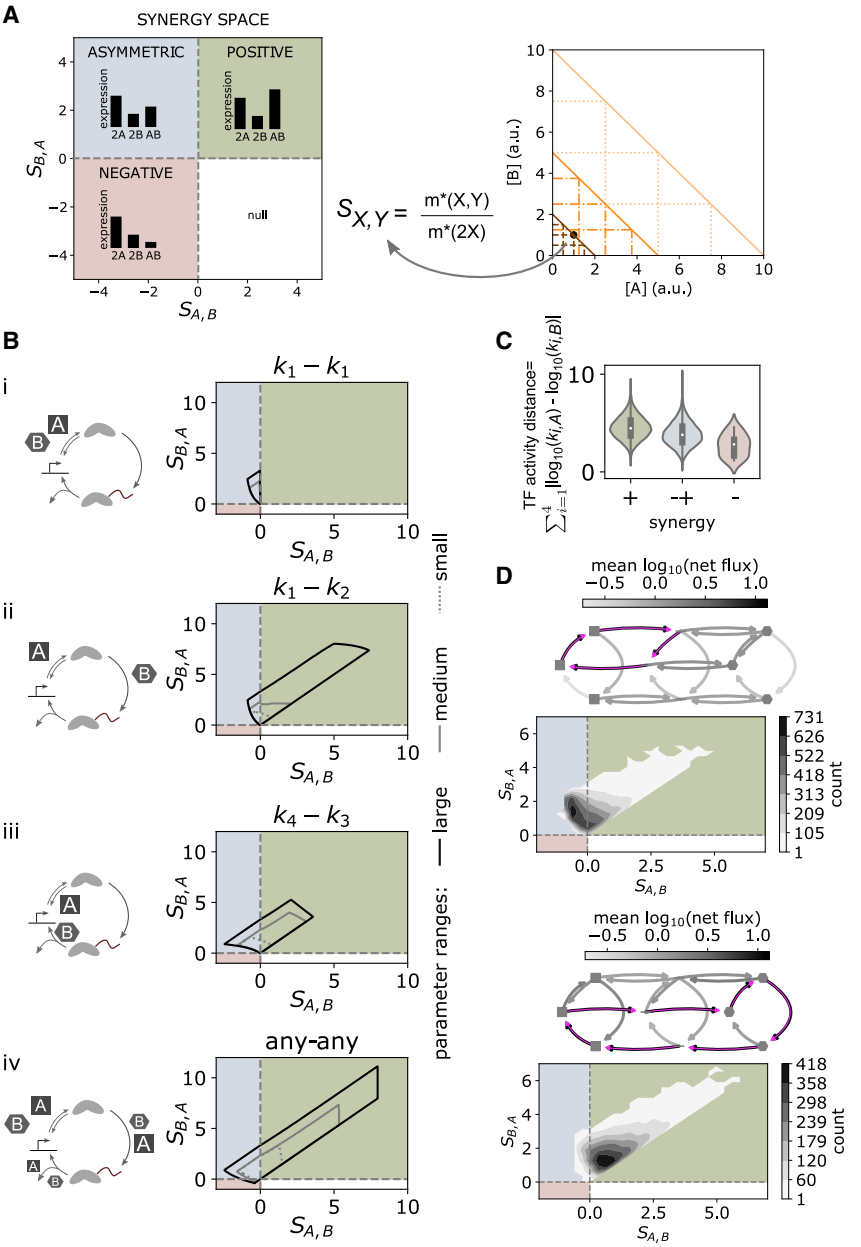


Figure 2. Positive, negative, or asymmetric synergy emerge in the model depending upon parameter values

(A) Left: synergy space. See text for details. Right: 2-dimensional space of TF concentrations where synergy could be defined. The axes represent the total concentrations of each individual TF. Synergy could be defined at any point in the diagonal that links the same two values for the totals, by comparing expression at the ratio given by that point, to expression at each of the individual totals on the diagonal extremes. The black dot indicates the 0.5:0.5: ratio at a total of two arbitrary units used in the work. The 0.75:0.25 and 0.25:0.75 ratios are also indicated by the broken lines.

(B) Regions of the synergy space spanned by 3 regulatory strategies (see also Figure S1A). (i) *A* and *B* act on the first step exclusively. (ii and iii) *A* on the first step (accelerates k_1 (ii) or reduces k_4 (iii)), *B* on a subsequent step. (iv) *A* and *B* act on any step (to various degrees). Constraints for the boundary search (STAR Methods, biologically plausible ranges for parameters): in all cases, binding and unbinding rates and TF-bound rates between 1 and 10^4. Basal cycle parameters: black: between 1 and 10^4, gray: between 1 and 100 for the clockwise rates, between 100–10,000 for $k_{4,\varnothing}$. TF effect on basal rates: black: at most 1,000 times larger than the basal rates for the clockwise or 0.001 times smaller for k_4 (1,000×/0.001×). Solid gray: 100×/0.01×. Dotted gray: 5×/0.2×. Fold change in m^* for each TF individually (at concentration 2) with respect to the basal condition with no TF: black: between 1 and 10; gray: between 1 and 5. The dotted gray line conditions are not plotted for k_1-k_1, because that region is already very small.

(C) Distribution of TF activity distances per synergy quadrant for a random sample of parameter sets under the same constraints as in (Biv) (corresponding synergies are plotted as a scatterplot in Figure S1Ci).

(D) The two most prevalent dominant flux paths for the points used in the analysis in (C). The arrow diagrams represent the model states and transitions, as schematized in Figure 1C; arrow grayscale intensity denotes the average probability net flux for that transition over all the parameter sets that share the dominant path highlighted in magenta. Note that reversible edges may appear in both directions if some parameter sets have net flux in one direction and others in the other. The distributions underneath show contours for the two-dimensional histogram of synergy values corresponding to those parameter sets that share the same dominant path. See also Figure S2.

measure at a total of 2 arbitrary units. Our analyses suggest that this ratio should maximize the possibility of detecting synergy. Yet, some of our results (Figure S6B) also suggest that a thorough exploration of these maps can potentially reveal further limits and mechanisms, and future systematic explorations may be useful.

Positive, negative, or asymmetric synergy can theoretically emerge from two activators

We begin by exploring the theoretically possible synergistic behaviors between two activators ($k_{1,X} \geq k_{1,\varnothing}$, $k_{2,X} \geq k_{2,\varnothing}$, $k_{3,X} \geq k_{3,\varnothing}$, $k_{4,X} \leq k_{4,\varnothing}$, $X \in \{A,B\}$) that have the same binding kinetics (given by a binding rate k_b and an unbinding rate k_u) but different kinetic roles (given by the $k_{i,X}$, $i \in \{1,2,3,4\}$, $X \in \{A,B\}$). We assume that the concentration unit is incorporated in the binding on-rates, such that both *A* and *B* are present at a concentration of 1 arbitrary unit each when they are both present together, and at concentration 2 when they are alone. In order to define the boundaries of the synergy space region that can be covered by the model under biologically plausible parameter values and constraints (STAR Methods, biologically plausible ranges for parameters), we numerically sampled the parameter space using a biased sampling algorithm (STAR Methods, synergy boundary for a regulatory strategy). We explored the synergy space

when TFs act on the same step, exclusively complementary steps, or to different extents on all steps (Figures 2B and S1A).

As a control, we first explored the case where both TFs enhance a single step. Figure 2Bi shows that as expected, only asymmetric synergy appears when both TFs accelerate k_1. The same result is obtained if both TFs accelerate k_2, k_3, or reduce k_4 (Figure S1A, second row). Intuitively, if TF *A* drives stronger expression than TF *B* but both act on the same step, then mixing *A* with *B* can only reduce expression with respect to the strongest one and increase it with respect to the weakest.

Next, we explored the case in which TFs have completely complementary kinetic roles and each of them acts on a different transition. Figures 2Bii–2Biii show the results when one TF acts on the first transition and the other acts on a subsequent transition. The plots show that this control strategy mostly results in positive synergy but also covers a region of the asymmetric synergy quadrant (notice that the result is restricted to the upper diagonal region of the positive quadrant because of the definition of TF *A* as the strongest of the pair). The appearance of asymmetric synergy in this case shows that even if TFs have complementary activities, it may not be enough to enhance expression beyond that of the strongest TF when half of its concentration is substituted by the weaker TF. A very similar result, with positive and asymmetric synergy, is obtained for any other pair of complementary rates except when one TF reduces k_4 and the other enhances k_2 in which case only asymmetric synergy emerges (Figure S1A). We provide an intuitive explanation for this behavior below.

In the two scenarios considered above, a given TF is considered to act only on one transition. However, this does not need to be the case. For example, if the first cycle transition is interpreted as the assembly of the pre-initiation complex (PIC) on DNA, the same activator TF can modulate PIC assembly by stabilizing its bound state, which would correspond to a reduction in k_4 but also by increasing the apparent binding on rate of various PIC components, which would correspond to an increase in k_1.[64,65] In addition, the 3-step cycle may also be interpreted more coarsely, with this first step accounting for chromatin modifications, which may also be mediated by TFs. Moreover, TFs are often found to interact with a wide range of cofactors and regulators,[66–68] contributing to the likelihood that they modulate multiple processes albeit with different strengths. Hence, we next considered a more general scenario where each TF can accelerate any of the clockwise transitions or reduce the backward transition to different extents (Figure 2Biv). In this case, a slightly higher region of the positive and asymmetric synergy quadrants is occupied, and slightly negative synergy can also emerge. We interpret this as an indication that under some parameter values, TFs can interfere with each other's action and reduce the expression as compared with when only one of them is present.

The black line in the plots of Figure 2B shows the results for the model exploration under biologically plausible parameter values that overall define very large parameter ranges. This allows us to define what kinds of behaviors are in theory possible out of a given regulatory strategy. However, it is well known that the effects of mammalian TFs are weak. Though accurate biochemical measurements are scarce, measurements in the literature suggest that a TF may change the rate of a given biochemical process by a 2–4-fold range[65,69] instead of the 1,000-fold range allowed in the analysis. The expression fold changes that we observed in our data (below) were in the 2–4-fold range as well instead of the 10-fold allowed in the exploration. When we constrain the model such that the basal transitions (in the absence of TF) cannot occur at the maximum rate, the TFs change a given basal rate up to a 100-fold or 5-fold (instead of a 1,000-fold), and the mRNA response fold change is also up to 5-fold (instead of 10-fold) the synergy space region that can be spanned by the model becomes much smaller (Figure 2B, gray lines). This suggests that for weak TFs operating through a single binding site, the accessible synergy region will be quite small.

The activity of the TFs over the cycle is not the only determinant of synergy

The original proposition of kinetic synergy stemmed from the assumption that synergy would emerge from TFs acting on different rate-limiting steps in transcription.[13] In the case of TFs with potentially overlapping effects, to what extent is positive synergy linked to TFs working exclusively, or nearly exclusively, on separate steps, so that they complement each other to enhance the cycle? In order to address this question, we looked at the correspondence between parameter values and synergy. For this, we generated a random sample of points that span a wide region of the synergy space (plotted in Figure S1Ci; STAR Methods, random sample of points in synergy space). In order to quantify the degree of complementarity between the pair of TFs in a given parameter set, we use the following measure, which we call TF activity distance: the sum, over all the polymerase cycle transitions, of the absolute differences between the logarithms of the transition rates associated to each TF (Figure 2C). Similar TF parameter values result in a small distance value, whereas TFs with big differences in their rates, and therefore more divergent in their functions, result in a larger distance. As shown in Figure 2C, positive synergy tends to emerge at higher distances than asymmetric and negative synergies, suggesting more divergent functions is indeed linked to higher complementarity and thus higher positive synergy.

However, the distances that lead to asymmetric synergy and those that lead to positive synergy overlap, suggesting that the different functions of the TFs are not the only determinants of synergy output. When binning the distributions by the basal expression (steady state m^* in the absence of TFs) and binding and unbinding rates, these factors appear to be important as well: higher basal expression and higher binding and unbinding rates correlate with less distant TFs producing positive synergy (Figure S1D). In addition, the basal expression and binding rates also modulate the correlation between the distance of two TFs and the extent of positive synergy that they exhibit (Figure S1E).

Intuitively, for positive synergy to emerge, we would expect that each of the TFs binds and unbinds appropriately as to be able to exert its effect and not interfere with the binding and the effect of the other TF. In order to test the extent to which this is indeed linked to synergy, we looked at the steady-state probability fluxes in the graph. Given the irreversible nature of the transitions of the polymerase cycle, a net probability flux remains even when the system is at steady state. The flux of probability of the system is intimately linked to the production of

mRNA, because mRNA is produced as the system transitions through the polymerase cycle. Formally, the flux from node i to node j, $J_{i,j}$ is given by $J_{i,j} = k_{i,j} P_i$, with $k_{i,j}$ the transition rate between i and j, and P_i the probability of node i. In the case of irreversible edges, this equals the net flux. In the case of reversible edges, the net flux $\overline{J_{i,j}}$ can be defined as $\overline{J_{i,j}} = J_{i,j} - J_{j,i}$, with $J_{i,j} > J_{j,i}$.

For the same sample of points (parameter sets) as in Figure 2C, we computed the net fluxes in the presence of A and B. Then, for each point, by starting at the polymerase-empty state with no TF bound (state $1,\varnothing$ in Figure 1B, $\mathcal{N} \times \mathcal{P}$) we followed the transition with a higher net flux, and repeated the same iteratively until reaching state $1,\varnothing$ again or any other node already encountered. This generates what we call the dominant path of net fluxes over the graph. After computing the dominant path for each of the parameter sets, we quantified how many parameter sets share the same dominant path. For this analysis, we pulled together those pairs of paths that are mirror images of each other because they are equivalent.

Out of all the parameter sets sampled, the majority correspond to one of two paths, represented in Figure 2D. The most predominant involves the binding of one TF, transition over the first step (binding of polymerase), unbinding of the TF, and reversion to the empty state. The two-dimensional density plot below the flux diagram shows that the majority of the points with this dominant path of fluxes correspond to asymmetric synergy. In contrast, the second most frequent dominant path involves cycling over the whole graph, with the first two transitions occurring under one TF, and the last occurring under the other. In this case, the majority of the points are associated with positive synergy. The rest of the dominant paths that make up to 90% of all the dominant paths in the sample of points are shown in Figure S2. The density plots show that dominant paths are not uniquely associated to individual synergy classes, but there are clear biases, with positive synergy being mostly associated to dominant paths that traverse the whole graph, and asymmetric synergy linked to dominant paths that show nonproductive cycling. This agrees with the expectation that positive synergy should emerge when TFs act productively to enhance progression over the polymerase cycle, but it also suggests that an intricate balance between all the transitions in the system is required for positive synergy to emerge.

This analysis also clarifies why no positive synergy emerges in the situation where one TF acts exclusively to reduce k_4 and the other acts exclusively to accelerate k_2 (Figure S1A, $k_4 - k_2$). In this case, both TFs need to be bound to state 2 of the cycle to exert their effect. However, as the TF that acts to slow down k_4 unbinds, the system transitions back to state 1 before the other TF can bind or act. This can be seen in the analysis of the steady-state fluxes shown in Figure S3A, for a random sample of points for this condition (Figure S1Cii). In contrast, full cycling is recovered when one TF slows down k_4 and the other accelerates k_3 (Figures S1Ciii and S3B).

It is important to note that because of the irreversibility of the cycle, cycling can also be enhanced when TFs have the same kinetic roles but different binding kinetics. As a result, if our original assumption that TFs have the same binding kinetics does not hold, positive synergy could emerge even if TFs do not interact functionally but bind differently, especially if TFs would bind to each cycle state with different kinetics (Figure S1B).

Overall, these modeling results show that this single binding site scenario and our synergy measure can reveal kinetic synergy. When two TFs bind to a shared site with the same kinetics, the observation of positive synergy indicates that TFs have at least partially complementary kinetic roles, which can enhance each other's effect.

Experimental evidence of kinetic synergy using a synthetic platform

In order to test whether functional interactions between TFs and corresponding kinetic synergy can be detected experimentally, we developed a reporter system in which synthetic TF fusions (synTFs) are recruited to a single DNA binding site integrated into a mammalian HEK293 cell line (STAR Methods, construct design and cloning and genomic integration of reporter constructs).[56–58] We engineered synTFs composed of an activation domain from a mammalian TF, fused to a synthetic zinc finger (ZF) DNA binding domain. This array of synthetic ZFs is designed to target a 20-bp binding site that does not natively exist in the mammalian genome sequence (Figures 1A and 3A).[56–58] This allows us to specifically recruit the activation domains to a reporter to assess their effects on transcription, while minimizing confounding effects from native TFs acting on the reporter. Moreover, because all TFs share the same binding domain, TFs should bind and unbind with very similar kinetics. Therefore, if positive synergy is observed, this should reflect functional interactions among the TFs.

We selected five activation domains of mammalian TFs with a described diversity of functions in the literature. SP1 is a ubiquitous mammalian TF whose mechanism of action has classically been linked to the recruitment of the transcriptional machinery.[70] cMyc is also a ubiquitous regulator. It interacts with a diverse range of proteins, but its mechanism of action has been predominantly linked to processes downstream of the recruitment of the transcriptional machinery, including pause-release[45] and elongation via interaction with the elongation factor Spt5.[46] BRD4 has also been described to have elongating activity, through the interaction with positive transcription elongation factor b (pTEF-b).[71,72] In addition, it has been involved in phase-separation at super-enhancers,[73] suggesting that BRD4 may also regulate other steps in the transcription cycle. Finally, we chose the activation domain of HSF1, which has been described to have both initiating and pause-release stimulating activity, and a mutant version of it, which we call HSF1-m. This mutant was described to be elongation-deficient.[74] Accordingly, these TFs can be broadly classified into either initiating (if they influence the recruitment of RNA polymerase) or elongating factors (if they influence a process downstream), as depicted in Figure 3A.

We then stably integrated into HEK293 cells a reporter, composed of a single target binding site upstream of a minimal CMV (minCMV) promoter driving the expression of a destabilized EGFP (d2EGFP) (STAR Methods, genomic integration of reporter constructs). Given its rapid turnover,[75] destabilized EGFP serves as a convenient genomic reporter of the mRNA expression level.[76] The expression of the synTFs was induced by transient transfection of the synTFs, whose expression level can be controlled by the amount of the plasmids transfected (Figure S4A). We chose to transfect synTFs at either 10 or 20 ng to ensure that the concentration (i.e., expression level) of synTF is

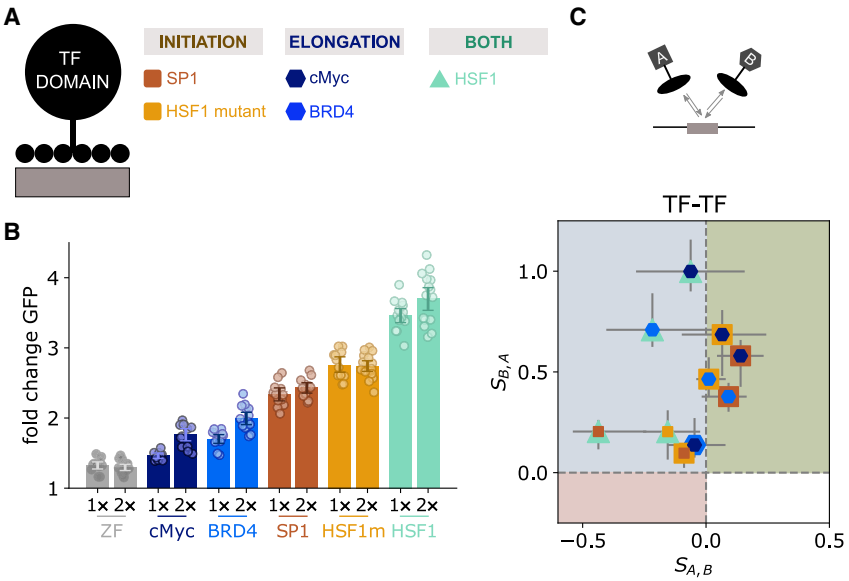


Figure 3. Experimental observation of kinetic synergy between 2 transcriptional activators on a synthetic platform

(A) Schema of the synthetic TFs.

(B) Expression upon transfection with either 10 (1×) or 20 (2×) ng of one TF, or the ZF binding domain alone (gray). Error bars denote the 95% confidence interval for the mean GFP fold change, obtained from bootstrapping the mean GFP fold-change values across all the measurements for each condition. At least 3 biological replicates per condition, with 2–4 technical replicates each. Each scatter point represents an individual measurement.

(C) Experimental synergy between two activators, defined as in equations 2 and 3 ($\log_2$ of the ratio of average fold-change expression when 10 ng of each TF is transfected, over the average fold-change expression when 20 ng of one is transfected). TF *A* is the strongest of the pair in the single TF expression, as shown in the 2× conditions of panel B. Error bars denote ranges from at least three biological replicates, with 2–4 technical replicates each. Individual synergy points for each biological replicate, as well as bar plots corresponding to the fold changes, are shown in Figures S4C and S4E. The synergy between each TF and the empty ZF is shown in Figure S4D.

the limiting factor. Reporter expression outcome was assessed by quantifying GFP fluorescence using flow cytometry 48 h later (STAR Methods, transient transfection and flow cytometry and data analysis).

Figure 3B shows reporter activation by each of the synTFs. We observed similar activation strengths varying from about 1.5-fold change in GFP fluorescence to 4-fold change, with slight increases upon doubling the amount of TF transfected for most TFs. Such fold change up-regulation is in the range of physiological induction in mammalian signaling pathways (e.g., Strasen et al.,[77] Wong et al.,[78] and Friedrich et al.[79]). A similar dose-dependent increase in reporter signal is also observed at the mRNA level (Figure S4B), supporting the use of GFP fluorescence as a readout.

In order to assess the extent of synergy between pairs of TFs, we compared the fold change in GFP fluorescence when TFs were transfected in pairs at 10 ng each with that when only one is transfected at 20 ng. We used quantitative immunofluorescence targeting the HA-tag of the synTFs to verify that transient transfection of 20 ng of coding plasmids for a single synTF results in a similar synTF abundance distribution as when transfecting two TFs in combination at 10 ng each, despite some variability inherent to the transfection procedure (Figure S5A) (STAR Methods, immunofluorescence to assess synergy for cMyc-SP1). Under these conditions, Figure 3C shows that both positive and asymmetric synergy appears. The error bars in the figure denote the ranges of the data, which is the synergy measures obtained from each of at least three independent biological replicates (See Figures S4C and S4E for details). We acknowledge that there is variability among experiments, because of the combination of small effects of the TFs and the variability associated with transfection and flow cytometry measurements. Yet, some pairs are consistent, in particular the positive synergy observed between cMyc and SP1 across all biological replicates.

Given the variability of transfection, and the small effects we observe experimentally, we performed immunofluorescence to simultaneously quantify total input synTF levels (HA-tag) and output GFP at the single-cell level. With these data, we can ensure that we are comparing GFP levels between cells in which the cells transfected with the two synTFs together do not have higher input levels than cells transfected with either synTF alone. With this experimental assay, we also found that the distribution of GFP fluorescence levels is shifted toward higher values when cells have been transfected with the combination of the two synTFs, thus confirming the positive synergy detected by flow cytometry (Figures S5C–S5E; STAR Methods, immunofluorescence to assess synergy for cMyc-SP1).

Consistent with the correlation in the model between TF activity distance and synergy class, the pairs exhibiting positive synergy (Figure 3C, green quadrant) correspond to those where each TF predominantly has been described to have either initiating or elongating factor activities. No TF was capable of increasing the expression from that driven by HSF1, which is the strongest synTF in the set and is described to have both initiating and elongating activities.[74] However, different TFs reduced its expression to different extents, suggesting some functional interactions are occurring (e.g., compare the $S_{A,B}$ coordinate for SP1-HSF1 and cMyc-HSF1 in Figure 3C). For the pairs of TFs described to predominantly act upon the same step, almost no synergy was detected (SP1-HSF1m, cMyc-BRD4).

Figure 3B shows a very modest activation effect from the ZF alone (no TF activation domain) case. However, the combination with a full synTF only leads to asymmetric synergy (Figure S4D), with the response to all TFs except HSF1 being reduced by the same extent, and that of HSF1 being reduced even further. This suggests that although the ZF may have a small effect, perhaps by increasing the ability of the basal transcriptional machinery to bind, the positive synergy observed between pairs of

TFs is most likely because of their activation domains, because the ZF only reduces expression when mixed with any of the TFs.

Overall, these results show that this single binding site scenario can be used to experimentally reveal kinetic synergy between TFs. As a result of using a single binding site and mammalian TFs, the individual activation effects are in the 2–4-fold range, and the synergistic effects are also small. This aligns with the modeling results when TFs are assumed to be weak (Figure 2B, gray lines). Moreover, the distributions of Figures S1D and S1E and the analysis of the dominant flux paths in Figure 2 point to binding and unbinding kinetics as important contributors to synergy as well. We now focus on this point.

Kinetic synergy depends upon the binding and unbinding kinetics

We explored how the synergy exhibited by a pair of TFs changes in the model as a function of either the unbinding or the binding rate. We began by examining the effect of the unbinding rate. To this end, we randomly sampled parameter sets for the basal rates over the polymerase cycle ($k_{1,\varnothing}$, $k_{2,\varnothing}$, $k_{3,\varnothing}$, $k_{4,\varnothing}$) and binding and unbinding (k_b, k_u). For each of these basal sets, we sampled parameter values for pairs of TFs ($k_{1,A}$, $k_{2,A}$, $k_{3,A}$, $k_{4,A}$, $k_{1,B}$, $k_{2,B}$, $k_{3,B}$, $k_{4,B}$). For each pair, we varied the unbinding rate k_u over a 2 order magnitude range, 10-fold up and down the basal value, and tracked the corresponding behavior over the synergy space. Given that the unbinding rate changes expression from each TF alone, we only considered those parameter sets where the strongest TF is the same across the unbinding rates considered, so that synergy is consistently defined throughout. Further details of this procedure are given in STAR Methods, effect of binding/unbinding rate on synergy.

To classify the behavior over the synergy space systematically, we considered that the binding and unbinding rate are related to affinity by $K_a = k_b[TF]/k_u$, and we used the relationship between changes in synergy and affinity so that the same criteria can be used to analyze the results when perturbing either the binding or the unbinding rate. We focused on the positive and asymmetric synergy behaviors and used a 4-bit string that captures the behavior at the affinity extremes: the first position denotes if $S_{A,B}$ is positive (p) or negative (n) at highest affinity, and the second position denotes the sign at the lowest affinity. The third and fourth positions denote whether $S_{A,B}$ and $S_{B,A}$ increase (i) or decrease (d), respectively, as affinity decreases. We disregard those situations where there is no change. As a result, there are theoretically 12 possible behaviors. We found that for some basal sets of parameters, changing the unbinding rate could result in all 12 possible behaviors, depending on the pair of TF parameter values. One such example is shown in Figure S6A, and selected examples are shown in Figure 4A. Similar results were found when modulating the binding on-rate k_b (Figure S6B), which can be interpreted as modulating the baseline concentration of the TFs at 1× concentration. This would be equivalent to moving along the 45° diagonal of the space illustrated in Figure 2A-right.

As expected from typical occupancy-based hypotheses, we found instances in which increasing affinity led to an increase in synergy (Figure 4A, more affinity-more synergy), changing from asymmetric to positive. In contrast, we also found examples in which even if the expression from the individual and combined TFs decreases with less affinity, synergy increases and can change from asymmetric to positive as affinity is reduced (Figure 4A, less affinity-more synergy). As seen in Figure S6 and depicted at the bottom of Figure 4A, we found many instances of non-monotonic behavior in which synergy was maximal at intermediate affinities.

To examine the relationship between the change in synergy class and the cycling over the system promoted by the TFs, we determined the dominant paths of net fluxes at steady state for parameter sets where synergy changes between asymmetric and positive or vice-versa as a function of the unbinding rate. We calculated the dominant path for the lowest and highest $S_{A,B}$ in the presence of both TFs. For each dominant path, we assessed whether it spanned nodes in each of the three binding configurations of the system (full path) or not (restricted path), as depicted for the corresponding examples of Figure 4A. Then, for each parameter set, we assessed whether the path type changed between the smallest and largest $S_{A,B}$ value and plotted the quantification in Figure 4B. As expected, and in line with the examples in Figure 4A, the barplot shows that in the majority of the cases, the change from smallest to highest $S_{A,B}$ value correlates with a transition from a restricted to a full dominant path. For the case where increasing the unbinding rate causes synergy to increase only with respect to TF *A* (npid), we found many instances with no change of path class and a small set where the relationship was reversed. This result aligns well with those of the previous sections, which show that the synergy of a pair of TFs ultimately depends on the overall system behavior and the intricate balance between all the transitions. However, a major contributor to the synergistic behavior of the TFs is the productive cycling over the system, with each TF binding and unbinding appropriately to allow the other to exert its effect.

Binding affinity variants of the synTFs can be generated by mutation to the DNA binding domain. A series of arginine residues in the ZF framework mediate nonspecific interactions with the DNA phosphate backbone.[80,81] Mutating these arginine residues to alanine reduces binding affinity in a dose-dependent manner.[30,56] To examine how changing binding affinity might affect the synergy between pairs of synTFs, we generated two affinity mutants for each synTF, with 5 or 7 mutations in the ZF domain (Figure S7). We refer to these constructs as 5× and 7× affinity mutants, or intermediate and lowest affinity, respectively. We quantified synergy for each of the synTF pairs at each of the two reduced affinities and combined the results with those from the WT affinity. The results for all pairs are shown in Figure S7.

In the majority of the cases, we found very small effects, with distinguishable average synergy values but overlapping ranges across conditions. However, in two pairs the data ranges were also distinguishable in the $S_{A,B}$ coordinate (* in Figure 4C, SP1-HSF1, HSF1m-HSF1), and there was an associated change in synergy class. We have also highlighted in Figure 4C other two pairs where there is a trend associated with a change in synergy class and an example with no effect. For some pairs, including those in which both TFs are thought to act on the same step, no distinguishable effects of affinity were observed (Figure 4C-no effect, Figure S7). Given the small effects and variability observed, we remain circumspect in our interpretation of this data. At a minimum, it suggests that the effects of affinity on

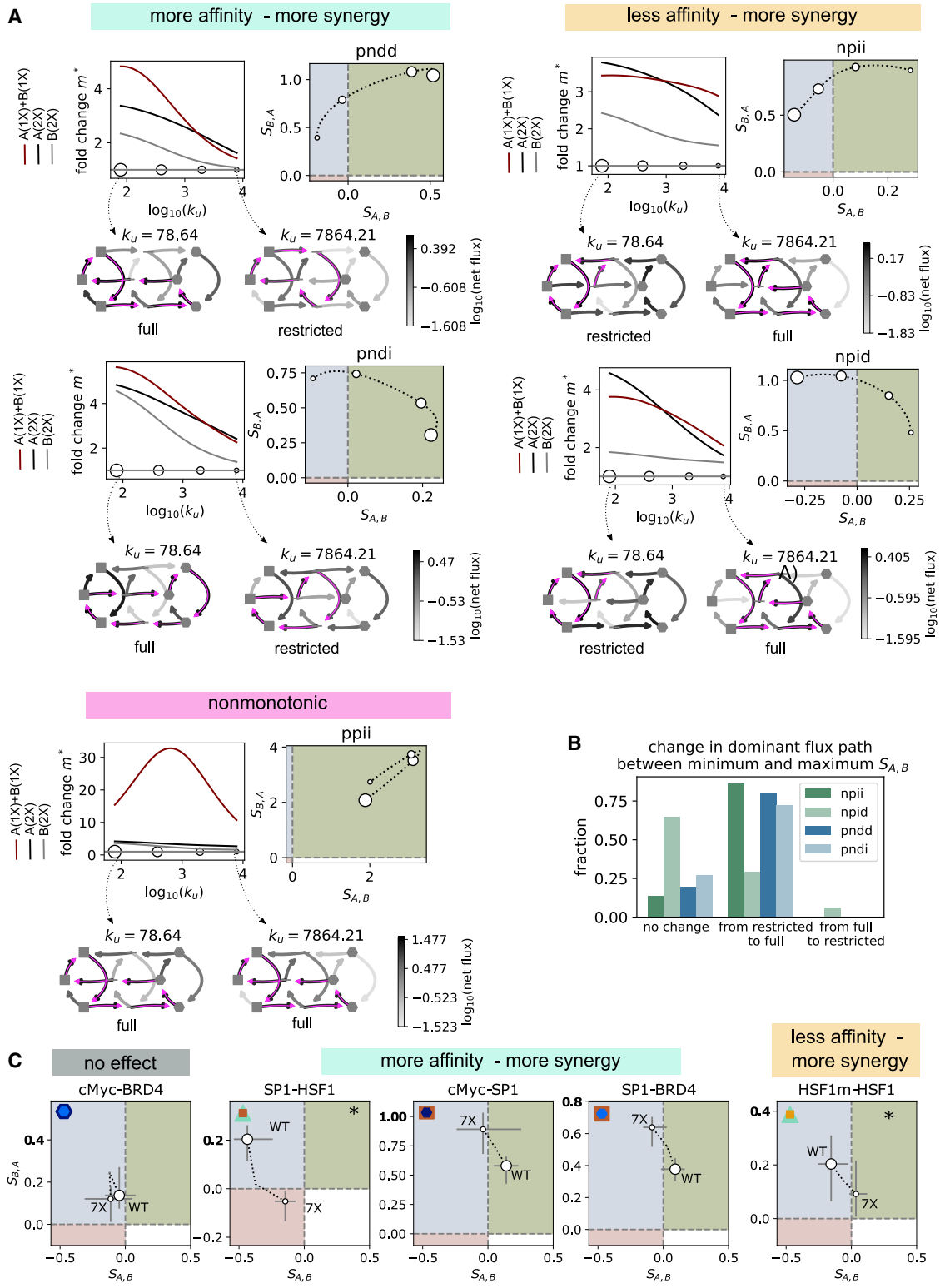


Figure 4. Synergy between a pair of TFs depends upon the binding and unbinding kinetics

(A) Model examples for 5 sets of parameter values demonstrating the diversity in how synergy changes as a function of the unbinding rate. For each example, the top-left plot shows the fold change in expression as compared with no TF present, for each of the TFs at concentration 2 (black, gray), or both TFs at concentration 1 (maroon), as a function of the unbinding rate. The top-right plot shows the corresponding behavior in synergy space. The circles on the bottom of the top-left plot and those on the top-right plot correspond to the same values of synergy. Marker size is related to binding affinity (smallest marker: smallest affinity, highest

(legend continued on next page)

synergy predicted by the model are experimentally tractable, and worth higher resolution and more in-depth interrogation in future studies.

DISCUSSION

In eukaryotic transcription, combinatorial control occurs at multiple scales, with many TFs binding to a given enhancer, and many enhancers controlling the activity of a gene.[24] Here, we have focused on the first scale, and have investigated how synergy between TFs can emerge as a result of the kinetics of the system. Although kinetic synergy was theoretically proposed almost 30 years ago,[13] its experimental demonstration has been challenging, largely because of the confound of cooperative binding interactions. To circumvent this limitation, we have focused on a scenario where only one TF can be specifically bound at any given time. By forcing the TFs to act separately in time while binding with the same kinetics, their functional interactions can be revealed. In order to reason about this scenario, we have proposed a minimal biophysical model that explicitly accounts for the kinetics of the binding, as well as the functional effects of the transcription factors over the transcription cycle. The model reveals that synergy between a pair of TFs is not an intrinsic feature of the pair but depends upon the balance between their binding and their functional effects. This work gives yet another example of the power of synthetic biology to answer fundamental biological questions.[30,63,82]

A model that explicitly accounts for the interplay between TF binding and polymerase activity

In order to reason about the single binding site experiment, we have developed a model with details of both the binding of the TFs and the progression over the polymerase cycle. This model brings together the two main modeling frameworks of transcription in the literature in which either the binding is taken implicitly (e.g., Scholes et al.[14]), or the polymerase cycle is not detailed (e.g., Estrada et al.[27]). In contrast to other attempts in the literature,[55] we do not make assumptions about the timescales of the binding and unbinding of the TFs with respect to those of the biochemical transitions over the polymerase cycle. This provides greater generality. In addition, the model can easily be extended to include more polymerase states and more binding sites for other TFs or coregulators, if such details become relevant in future studies. One of the simplifying assumptions of the model is that TFs only exert their effect while they are bound. We note that this does not necessarily have to be the case, because they may act through other cofactors that can remain bound even if the TF unbinds. This could be easily incorporated at the expense of more states and parameters. However, we think it would not fundamentally change our conclusions, because there would also be an interplay between the binding kinetics of these other components and the kinetic effects on the cycle.

We have explored the behavior of the model in parameter space under the assumption that the system is at steady state. This is a widely used assumption and reasonable for our experimental setup, given the time between transfection and measurement of mRNA levels. However, one of the contexts in which combinatorial control is most relevant is development, and many developmental processes may be too fast to allow for a steady state to be reached. In this case, it may become important to explicitly incorporate the time delay that emerges from polymerase traveling along the gene body, which we have not accounted for. Although at steady state this is likely to be effectively incorporated by the parameter of the last transition rate in the polymerase cycle, it could have important implications when considering how synergy emerges in transient regimes, and it will be a relevant point to consider in future studies.

Measuring synergy theoretically and experimentally

In order to quantify synergy, it has been common to measure the deviation from additivity, under the assumption that if TFs do not interact, then their combined effect should be the sum of the effects obtained when each TF is present alone.[16] Multiplicativity has also been taken as a measure of synergy.[15] However, we showed that when TFs interact functionally on a 2-step cycle, additivity or multiplicativity is only expected under very restricted circumstances.[14] As a result, synergy must be defined carefully under each model and experimental system, as recently pointed out for interactions between signaling pathway ligands.[83] Exploiting our single binding site synthetic system, here we compare expression when both TFs are present together with expression when only one of them is present, under the same total TF concentration. If TFs bind with the same kinetics and are at the same total concentration, changes in the response when there are two TFs as compared with only one must be because of their functional interactions. Our modeling analysis of Figures 2 and S1 show that, indeed, positive synergy only arises when TFs are allowed to act on partially or completely complementary transitions. In line with the original hypothesis for kinetic synergy, our analysis of the steady state fluxes shows that this is because of TFs productively enhancing the polymerase cycle when acting

unbinding rate). Shown below are the diagrams depicting the net fluxes (gray colormap) and dominant flux path (magenta) for the two extreme k_u values. All examples share the same basal parameter values: $k_{1,\varnothing} = 4.288, k_{2,\varnothing} = 11.023, k_{3,\varnothing} = 3.414, k_{4,\varnothing} = 10.362. k_b = 180.19$. TF associated parameter values are as follows: pndd : $k_{1,A} = 120.985, k_{2,A} = 154.358, k_{3,A} = 4.561, k_{4,A} = 2.854, k_{1,B} = 5.007, k_{2,B} = 25.685, k_{3,B} = 15.086, k_{4,B} = 2.083$; pndi : $k_{1,A} = 6.317, k_{2,A} = 517.659, k_{3,A} = 1433.877, k_{4,A} = 1.095, k_{1,B} = 11.275, k_{2,B} = 326.127, k_{3,B} = 15.328, k_{4,B} = 10.223$; npii : $k_{1,A} = 4.844, k_{2,A} = 6345.641, k_{3,A} = 151.500, k_{4,A} = 7.354, k_{1,B} = 4.504, k_{2,B} = 17.664, k_{3,B} = 2601.429, k_{4,B} = 3.088$; npid : $k_{1,A} = 6.784, k_{2,A} = 740.850, k_{3,A} = 56.436, k_{4,A} = 2.010, k_{1,B} = 4.821, k_{2,B} = 11.997, k_{3,B} = 909.506, k_{4,B} = 8.354$; ppii : $k_{1,A} = 937.265, k_{2,A} = 8084.904, k_{3,A} = 5.392, k_{4,A} = 1.982, k_{1,B} = 9.945, k_{2,B} = 18.372, k_{3,B} = 2047.513, k_{4,B} = 8.447$; See also Figure S6.

(B) Quantification of the change in dominant path in the presence of both TFs, from the smallest to the largest $S_{A,B}$. The parameter values were obtained from a rejection-based sampling algorithm, as explained in STAR Methods, effect of binding/unbinding rate on synergy. The number of parameter sets analyzed for each class are as follows: npii: 13,103 parameter sets, corresponding to 214 basal parameter sets. npid: 4,461, corresponding to 264 basal parameter sets. pndd: 2,833, corresponding to 87 basal parameter sets. pndi: 2,215, corresponding to 132 basal parameter sets.

(C) Experimental results showing the effects of changing synTF affinity on synergy for selected synTF pairs. White markers: mean (WT, 7×). Error bars: ranges of the biological repeats. Dotted line: line that links the means for the WT, 5× (marker and errors not shown), and 7× affinity pairs. (*) denotes the pairs where the means in $S_{A,B}$ values differ by at least 0.1 between the highest and lowest affinities, and the biological replicates are non-overlapping. See also Figure S7.

in combination, by binding and unbinding appropriately to allow each TF to exert its effect.

In addition to positive synergy, we define asymmetric and negative synergy. This enables the quantitative characterization of the synergy between a pair of TFs as a function of a variable of the system, by looking at the corresponding trajectory in synergy space. Moreover, we could evaluate synergy as a function of TF ratio, which might reveal further insight about TF function.

Our previous analyses had suggested that assessing synergy might be a way to elucidate the mechanism of action of TFs.[14] However, the current analysis shows that this is confounded by the effect of the binding kinetics. Moreover, parameter constraints that generate positive synergy in the model also generate asymmetric synergy. In this case, even if TFs may have complementary activities, their binding patterns may be imbalanced and may not allow productive interaction. In the case where either of the TFs works exclusively on one of two complementary steps, this contrasts with the finding of exclusively greater-than-additive behavior by,[14] highlighting the importance of accounting for the binding kinetics. Conversely, if TFs have the same kinetic role but do not bind with the exact same kinetics, the model shows that positive synergy could also arise just as a result of the different binding (Figure S1B). In our experimental system, in principle, we expect that TFs that share the same binding domain bind with very similar kinetics. However, we are aware of examples in the literature where the activation domain also affects binding,[84] and this could be further investigated in the future.

Kinetic synergy can experimentally emerge when two TFs time-share a binding site

The model shows that positive synergy is theoretically possible when TFs share a binding site, and synergy between a pair of TFs is strongly influenced by their binding kinetics. Theoretically, both the binding on-rate and off-rate can modulate the synergy exhibited by a pair of TFs, and lower affinity can either decrease or increase their synergy. In some cases, the compromise is evidenced as a non-monotonic effect of affinity upon synergy.

Our experimental data are consistent with these findings: in line with the concept that kinetic synergy arises from complementarity in the function of TFs, positive synergy is only detected for pairs where TFs are described to act predominantly on different steps in transcription (Figure 3C). Also, in line with the model, we find that synergy changes with affinity in both increased and decreased directions (Figure 4). We acknowledge that our experimental effects are small. On the one hand, this is consistent with the known small effects of mammalian TFs acting from a single site and the reduced synergy space in the model under the assumption that TFs are weak (Figure 2B). In addition, we are aware that there are some technical limitations to our approach that may obscure some of the effects. In the experiments, the TF concentration is controlled by transfection, which is intrinsically noisy, and the output is measured by flow cytometry, which is a coarse read-out. Combined with the small effects of the TFs and the affinity mutations, the biological effects might be masked or confounded in some cases. Given these limitations, we exploited the medium throughput of the assay to scan multiple TF pairs and focused on the consistency across biological replicates to support the modeling results.

In addition, we confirmed the positive synergy observed for the most consistent pair using immunofluorescence, which enables quantifying input-output in individual cells, which controls for variability in the transfection and flow cytometry assay. In the future, the experimental data can be further refined by carefully modulating the input concentration with an inducible system, and single-cell quantification of the output expression. Ideally, live-cell imaging should also help illuminate the kinetic aspects of kinetic synergy.

Implications for gene regulation in natural scenarios

Though the effects of synergy that we measure are small, they may nonetheless be significant for gene regulation. Small-effect risk variants can underlie disease (e.g., Manolio et al.[85]), and small-effect variations are proposed to underlie evolution (e.g., Gerhart et al.[86]). We propose that the functional complementarity between TFs, even if the effects are small, is likely to also be a significant factor to consider when reasoning about gene regulation by combinations of TFs.

In endogenous enhancers, some TFs do have overlapping binding sites as in our setup.[87–89] However, most typically, each TF has its own binding site. Even in this case, binding kinetics may still be important. The residence time of the TF on the DNA must be long enough for it to be able to exert its function. However, it is plausible that there could be interference either directly or through recruited cofactors, such that output may be maximized at intermediate affinities. This could be another reason behind the widespread presence of relatively low-affinity binding sites in eukaryotes[90–93] and the observation of fast TF binding kinetics.[94–96] Moreover, tuning binding site affinity might be an effective way to modulate expression beyond fully adding or removing a binding site, which could have evolutionary implications.[97] Along the same lines, kinetic synergy relaxes the need for strict arrangements between binding sites, another typical feature of eukaryotic transcriptional control.[4,98,99]

TF activity has often been considered to be modular. In this view, the activity of the activation domain is independent of that of the binding domain, which is assumed to be important only to target the TF to specific sites on the genome.[100] Evidence against this model includes allosteric interactions between the DNA binding domain and the activation domain,[101] and the observation that the activation domain may be involved in DNA recognition.[84] Adding to this, our work highlights the importance of considering TFs as a unit, where the binding and activation domains together dictate the effect of the TF. Our study emphasizes the value of considering an integrated view of transcriptional control, where the effect of a TF has to be understood in terms of the other components of the system.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability

- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture
- METHOD DETAILS
  - Modelling details and the linear framework
  - Biologically plausible ranges for parameters
  - Synergy boundary for a regulatory strategy
  - Random sample of points in synergy space
  - Effect of binding/unbinding rate on synergy
  - Construct design and cloning
  - Genomic integration of reporter constructs
  - Transient transfection
  - Flow cytometry and data analysis
  - Western blotting
- QUANTITATIVE REAL-TIME PCR
  - Immunofluorescence to assess synergy for cMyc-SP1

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2023.02.003.

ACKNOWLEDGMENTS

We thank members of the DePace, Khalil, and Gunawardena labs for helpful discussions. We thank Mike Springer for early conversations that inspired the single binding site experiment. Fluorescence microscopy experiments were performed at the Nikon Imaging Center at Harvard Medical School. We thank Jennifer Waters and Rylie Walsh for guidance and support. The computations in this paper were run on the O2 cluster supported by the Harvard Medical School Research Computing Group. This work was supported by NSF grant MCB-1713855 (A.S.K.)/ MCB-1715184 (A.H.D.), NIH grant 1R01GM 122928 (A.H.D. and J.G.), NIH grant R01EB029483 (A.S.K.), NSF CAREER IOS-1452557 (A.H.D.), EMBO Fellowship 683–2019 (R.M.C.), and the Giovanni Armenise Harvard Foundation (A.H.D.). A.S.K. also acknowledges funding from the NIH Director's New Innovator award (1DP2AI131083-01), a DoD Vannevar Bush Faculty Fellowship (N00014-20-1-2825), and a DARPA Young Faculty award (D16AP00142). A.H.D. also gratefully acknowledges funding from the Lynch Foundation Systems Biology Fellowship, the McKenzie Family Charitable Trust Systems Biology Fellowship Fund, and the John and Virginia Kaneb Fellowship award at Harvard Medical School.

AUTHOR CONTRIBUTIONS

Conceptualization, R.M.-C., M.P., K.B., C.S., A.S.K., J.G., and A.H.D.P.; methodology, R.M.-C., M.P., K.B., C.S., and J.G.; investigation, M.P., K.B., D.F., and C.S.; formal analysis, R.M.-C.; software, R.M.-C. and D.F.; validation, D.F.; writing (original draft), R.M.-C.; writing (review and editing), R.M.-C., M.P., K.B., D.F., A.S.K., J.G., and A.H.D.P.; resources, A.S.K., funding, A.S.K., J.G., and A.H.D.P.; supervision, A.S.K. and A.H.D.P.

DECLARATION OF INTERESTS

A.S.K. and M.P. are inventors on patents related to the ZF technology. A.S.K. is a scientific advisor for and holds equity in Senti Biosciences and Chroma Medicine and is a co-founder of Fynch Biosciences and K2 Biotechnologies. A.H.D. is a member of the advisory board of Cell Systems.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: November 30, 2021
Revised: August 22, 2022
Accepted: February 10, 2023
Published: April 19, 2023

REFERENCES

1. Field, A., and Adelman, K. (2020). Evaluating enhancer function and transcription. Annu. Rev. Biochem. *89*, 213–234.
2. Wunderlich, Z., and Mirny, L.A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. Trends Genet. *25*, 434–440.
3. Georges, A.B., Benayoun, B.A., Caburet, S., and Veitia, R.A. (2010). Generic binding sites, generic DNA-binding domains: where does specific promoter recognition come from? FASEB J. *24*, 346–356.
4. Smith, R.P., Taher, L., Patwardhan, R.P., Kim, M.J., Inoue, F., Shendure, J., Ovcharenko, I., and Ahituv, N. (2013). Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. Nat. Genet. *45*, 1021–1028.
5. Vandel, J., Cassan, O., Lèbre, S., Lecellier, C.H., and Bréhélin, L. (2019). Probing transcription factor combinatorics in different promoter classes and in enhancers. BMC Genomics *20*, 1–19.
6. Inukai, S., Kock, K.H., and Bulyk, M.L. (2017). Transcription factor–DNA binding: beyond binding site motifs. Curr. Opin. Genet. Dev. *43*, 110–119.
7. Ouyang, Z., Zhou, Q., and Wong, W.H. (2009). ChIP-seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. Proc. Natl. Acad. Sci. USA *106*, 21521–21526.
8. de Boer, C.G., Vaishnav, E.D., Sadeh, R., Abeyta, E.L., Friedman, N., and Regev, A. (2020). Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. Nat. Biotechnol. *38*, 56–65.
9. Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr. Opin. Genet. Dev. *43*, 73–81.
10. King, D.M., Hong, C.K.Y., Shepherdson, J.L., Granas, D.M., Maricque, B.B., and Cohen, B.A. (2020). Synthetic and genomic regulatory elements reveal aspects of cis-regulatory grammar in mouse embryonic stem cells. eLife *9*, 1–24.
11. Nie, Z., Guo, C., Das, S.K., Chow, C.C., Batchelor, E., Simons, S.S., and Levens, D. (2020). Dissecting transcriptional amplification by MYC. eLife *9*, 1–32.
12. Carey, M., Lin, Y.-s., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature *345*, 361–364.
13. Herschlag, D., and Johnson, F.B. (1993). Synergism in transcriptional activation: a kinetic view. Genes Dev. *7*, 173–179.
14. Scholes, C., DePace, A.H., and Sánchez, Á. (2017). Combinatorial gene regulation through kinetic control of the transcription cycle. Cell Syst. *4*, 97–108.e9.
15. Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., and Phillips, R. (2005a). Transcriptional regulation by the numbers: applications. Curr. Opin. Genet. Dev. *15*, 125–135.
16. Carey, M. (1998). The enhanceosome and transcriptional synergy. Cell *92*, 5–8.
17. Veitia, R.A. (2003). A sigmoidal transcriptional response: cooperativity, synergy and dosage effects. Biol. Rev. Camb. Philos. Soc. *78*, 149–170.
18. Michida, H., Imoto, H., Shinohara, H., Yumoto, N., Seki, M., Umeda, M., Hayashi, T., Nikaido, I., Kasukawa, T., Suzuki, Y., and Okada-Hatakeyama, M. (2020). The number of transcription factors at an enhancer determines switch-like gene expression. Cell Rep. *31*, 107724.
19. Ptashne, M. (2005). Regulation of transcription: from lambda to eukaryotes. Trends Biochem. Sci. *30*, 275–279.
20. Ackers, G.K., Johnson, A.D., and Shea, M.A. (1982). Quantitative model for gene regulation by lambda phage repressor. Proc. Natl. Acad. Sci. USA *79*, 1129–1133.
21. Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005b). Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev. *15*, 116–124.
22. Vashee, S., Willie, J., and Kodadek, T. (1998). Synergistic activation of transcription by physiologically unrelated transcription factors through

cooperative DNA-Binding. Biochem. Biophys. Res. Commun. *247*, 530–535.

23. Ambrosetti, D.C., Schöler, H.R., Dailey, L., and Basilico, C. (2000). Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. J. Biol. Chem. *275*, 23387–23397.
24. Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.
25. Frank, T.D., Carmody, A.M., and Kholodenko, B.N. (2012). Versatility of cooperative transcriptional activation: a thermodynamical modeling analysis for greater-than-additive and less-than-additive effects. PLoS One *7*, e34439.
26. Goldstein, I., Paakinaho, V., Baek, S., Sung, M.-H., and Hager, G.L. (2017). Synergistic gene expression during the acute phase response is characterized by transcription factor assisted loading. Nat. Commun. *8*, 1849.
27. Estrada, J., Wong, F., DePace, A., and Gunawardena, J. (2016). Information integration and energy expenditure in gene regulation. Cell *166*, 234–244.
28. Panne, D., Maniatis, T., and Harrison, S.C. (2007). An atomic model of the interferon-beta enhanceosome. Cell *129*, 1111–1123.
29. Malik, S., and Roeder, R.G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nat. Rev. Genet. *11*, 761–772.
30. Bashor, C.J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J.J., and Khalil, A.S. (2019). Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. Science *364*, 593–597.
31. Biddle, J.W., Martinez-Corral, R., Wong, F., and Gunawardena, J. (2021). Allosteric conformational ensembles have unlimited capacity for integrating information. eLife *10*, 1–58.
32. Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. Proc. Natl. Acad. Sci. USA *107*, 22534–22539.
33. Narasimhan, K., Pillay, S., Huang, Y.H., Jayabal, S., Udayasuryan, B., Veerapandian, V., Kolatkar, P., Cojocaru, V., Pervushin, K., and Jauch, R. (2015). DNA-mediated cooperativity facilitates the co-selection of cryptic enhancer sequences by SOX2 and PAX6 transcription factors. Nucleic Acids Res. *43*, 1513–1528.
34. Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186–192.
35. Mao, C., Brown, C.R., Falkovskaia, E., Dong, S., Hrabeta-Robinson, E., Wenger, L., and Boeger, H. (2010). Quantitative analysis of the transcription control mechanism. Mol. Syst. Biol. *6*, 1–12.
36. Hansen, A.S., and O'Shea, E.K. (2013). Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. Mol. Syst. Biol. *9*, 704.
37. Cui, G., Dong, Q., Duan, J., Zhang, C., Liu, X., and He, Q. (2020). NC2 complex is a key factor for the activation of catalase-3 transcription by regulating H2a.Z deposition. Nucleic Acids Res. *48*, 8332–8348.
38. Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol. *16*, 167–177.
39. Core, L., and Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. Genes Dev. *33*, 1–3.
40. Oven, I., Brdičková, N., Kohoutek, J., Vaupotič, T., Narat, M., and Peterlin, B.M. (2007). AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells. Mol. Cell. Biol. *27*, 8815–8823.
41. Wong, F., and Gunawardena, J. (2020). Gene regulation in and out of equilibrium. Annu. Rev. Biophys. *49*, 199–226.
42. Harden, T.T., Vincent, B.J., and DePace, A.H. (2023). Transcriptional activators in the early *Drosophila* embryo perform different kinetic roles. Cell Systems *14*. https://doi.org/10.1101/2021.02.25.432925.
43. Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996). Three functional classes of transcriptional activation domain. Mol. Cell. Biol. *16*, 2044–2055.
44. Fu, D., Wen, Y., and Ma, J. (2004). The co-activator CREB-binding protein participates in enhancer-dependent activities of bicoid. J. Biol. Chem. *279*, 48725–48733.
45. Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). C-Myc regulates transcriptional pause release. Cell *141*, 432–445.
46. Baluapuri, A., Hofstetter, J., Dudvarski Stankovic, N., Endres, T., Bhandare, P., Vos, S.M., Adhikari, B., Schwarz, J.D., Narain, A., Vogt, M., et al. (2019). MYC recruits SPT5 to RNA polymerase II to promote processive transcription elongation. Mol. Cell *74*. 674.e11–687.e11.
47. Bell, C.C., Talarmain, L., Scolamiero, L., Lam, E.Y.N., Ang, C.-S., Gilan, O., and Dawson, M.A. (2022). Transcription factors use a unique combination of cofactors to potentiate different promoter-dependent steps in transcription. Preprint at bioRxiv. https://doi.org/10.1101/2022.10.25.513774.
48. Danko, C.G., Hah, N., Luo, X., Martins, A.L., Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol. Cell *50*, 212–222.
49. Vanhille, L., Griffon, A., Maqbool, M.A., Zacarias-Cabeza, J., Dao, L.T.M., Fernandez, N., Ballester, B., Andrau, J.C., and Spicuglia, S. (2015). High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. Nat. Commun. *6*, 6905.
50. Singh, G., Mullany, S., Moorthy, S.D., Zhang, R., Mehdi, T., Tian, R., Duncan, A.G., Moses, A.M., and Mitchell, J.A. (2021). A flexible repertoire of transcription factor binding sites and a diversity threshold determines enhancer activity in embryonic stem cells. Genome Res. *31*, 564–575.
51. Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., and Khalil, A.S. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell *158*, 110–120.
52. Stampfel, G., Kazmar, T., Frank, O., Wienerroither, S., Reiter, F., and Stark, A. (2015). Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature *528*, 147–151.
53. Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. Science *332*, 472–474.
54. Rybakova, K.N., Bruggeman, F.J., Tomaszewska, A., Moné, M.J., Carlberg, C., and Westerhoff, H.V. (2015). Multiplex eukaryotic transcription (in)activation: timing, bursting and cycling of a ratchet clock mechanism. PLoS Comput. Biol. *11*, 1–24.
55. Li, C., Cesbron, F., Oehler, M., Brunner, M., and Höfer, T. (2018). Frequency modulation of transcriptional bursting enables sensitive and rapid gene regulation. Cell Syst. *6*, 409–423.e11.
56. Khalil, A.S., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, J.K., and Collins, J.J. (2012). A synthetic biology framework for programming eukaryotic transcription functions. Cell *150*, 647–658.
57. Park, M., Patel, N., Keung, A.J., and Khalil, A.S. (2019b). Engineering epigenetic regulation using synthetic read-write modules. Cell *176*. 227.e20–238.e20.
58. Israni, D.V., Li, H.-S., Gagnon, K.A., Sander, J.D., Roybal, K.T., Joung, J.K., Wong, W.W., and Khalil, A.S. (2021). Clinically-driven design of synthetic gene regulatory programs in human cells. Preprint at bioRxiv. https://doi.org/10.1101/2021.02.22.432371.
59. Gunawardena, J. (2012). A linear framework for time-scale separation in nonlinear biochemical systems. PLoS One *7*, e36321.
60. Ahsendorf, T., Wong, F., Eils, R., and Gunawardena, J. (2014). A framework for modelling gene regulation which accommodates non-equilibrium mechanisms. BMC Biol. *12*, 102.
61. Nam, K.-M., Martinez-Corral, R., and Gunawardena, J. (2022). The linear framework: using graph theory to reveal the algebra and thermodynamics of biomolecular systems. Interface Focus *12*, 20220013.

62. Biddle, J.W., Nguyen, M., and Gunawardena, J. (2019). Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA. eLife *8*, 1–45.

63. Park, J., Estrada, J., Johnson, G., Vincent, B.J., Ricci-Tam, C., Bragdon, M.D., Shulgina, Y., Cha, A., Wunderlich, Z., Gunawardena, J., et al. (2019a). Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. eLife *8*, e41266.

64. Baek, I., Friedman, L.J., Gelles, J., and Buratowski, S. (2021). Single-molecule studies reveal branched pathways for activator-dependent assembly of RNA polymerase II pre-initiation complexes. Mol. Cell *81*. 3576.e6–3588.e6.

65. Coleman, R.A., Qiao, Z., Singh, S.K., Peng, C.S., Cianfrocco, M., Zhang, Z., Piasecka, A., Aldeborgh, H., Basishvili, G., and Liu, W.L. (2017). p53 dynamically directs TFIID assembly on target gene promoters. Mol. Cell. Biol. *37*, e00085-17.

66. Dingar, D., Kalkat, M., Chan, P.K., Srikumar, T., Bailey, S.D., Tu, W.B., Coyaud, E., Ponzielli, R., Kolyar, M., Jurisica, I., et al. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. J. Proteomics *118*, 95–111.

67. Kim, B.R., Coyaud, E., Laurent, E.M.N., St-Germain, J., Van De Laar, E., Tsao, M.S., Raught, B., and Moghal, N. (2017). Identification of the SOX2 interactome by BioID reveals EP300 as a mediator of SOX2-dependent squamous differentiation and lung squamous cell carcinoma growth. Mol. Cell. Proteomics *16*, 1864–1888.

68. Carnesecchi, J., Sigismondo, G., Domsch, K., Baader, C.E.P., Rafiee, M.R., Krijgsveld, J., and Lohmann, I. (2020). Multi-level and lineage-specific interactomes of the Hox transcription factor Ubx contribute to its functional specificity. Nat. Commun. *11*, 1388.

69. Liu, J., Hansen, D., Eck, E., Kim, Y.J., Turner, M., Alamos, S., and Garcia, H.G. (2021). Real-time single-cell characterization of the eukaryotic transcription cycle reveals correlations between RNA initiation, elongation, and cleavage. PLoS Comput. Biol. *17*, e1008999.

70. O'Connor, L., Gilmour, J., and Bonifer, C. (2016). The role of the ubiquitously expressed transcription factor Sp1 in tissue-specific transcriptional regulation and in disease. Yale J. Biol. Med. *89*, 513–525.

71. Yang, Z., Yik, J.H.N., Chen, R., He, N., Jang, M.K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Mol. Cell *19*, 535–545.

72. Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol. Cell *19*, 523–534.

73. Vasile, E., Hnisz, D., Klein, I.A., Young, R.A., Manteiga, J.C., Malik, S., Lee, T.I., Abraham, B.J., Schuijers, J., Cisse, I.I., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

74. Brown, S.A., Weirich, C.S., Newton, E.M., and Kingston, R.E. (1998). Transcriptional activation domains stimulate initiation and elongation at different times and via different residues. EMBO J. *17*, 3146–3154.

75. Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., and Kain, S.R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. J. Biol. Chem. *273*, 34970–34975.

76. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS Biol. *4*, 1707–1719.

77. Strasen, J., Sarma, U., Jentsch, M., Bohn, S., Sheng, C., Horbelt, D., Knaus, P., Legewie, S., and Loewer, A. (2018). Cell-specific responses to the cytokine TGF β are determined by variability in protein levels. Mol. Syst. Biol. *14*, 1–17.

78. Wong, V.C., Mathew, S., Ramji, R., Gaudet, S., and Miller-Jensen, K. (2019). Fold-change detection of NF-κB at target genes with different transcript outputs. Biophys. J. *116*, 709–724.

79. Friedrich, D., Friedel, L., Finzel, A., Herrmann, A., Preibisch, S., and Loewer, A. (2019). Stochastic transcription in the p53-mediated response to DNA damage is modulated by burst frequency. Mol. Syst. Biol. *15*, 1–20.

80. Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809–817.

81. Elrod-Erickson, M., Rould, M.A., Nekludova, L., and Pabo, C.O. (1996). Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. Structure *4*, 1171–1180.

82. Crocker, J., Tsai, A., and Stern, D.L. (2017). A fully synthetic transcriptional platform for a multicellular eukaryote. Cell Rep. *18*, 287–296.

83. Klumpe, H.E., Langley, M.A., Linton, J.M., Su, C.J., Antebi, Y.E., and Elowitz, M.B. (2022). The context-dependent, combinatorial logic of BMP signaling. Cell Systems *13*, 388–407.e10.

84. Brodsky, S., Jana, T., Mittelman, K., Chapal, M., Kumar, D.K., Carmi, M., and Barkai, N. (2020). Intrinsically disordered regions direct transcription factor in vivo binding specificity. Mol. Cell *79*, 1–3.

85. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747–753.

86. Gerhart, J., and Kirschner, M. (2007). The theory of facilitated variation. Proc. Natl. Acad. Sci. USA *104* (*Suppl 1*), 8582–8589.

87. Han, W., Yu, Y., Su, K., Kohanski, R.A., and Pick, L. (1998). A binding site for multiple transcriptional activators in the fushi tarazu proximal enhancer is essential for gene expression in vivo. Mol. Cell. Biol. *18*, 3384–3394.

88. Pan, Y., and Nussinov, R. (2011). The role of response elements organization in transcription factor selectivity: the IFN-β enhanceosome example. PLoS Comput. Biol. *7*, e1002077.

89. Cheng, Q., Kazemian, M., Pham, H., Blatti, C., Celniker, S.E., Wolfe, S.A., Brodsky, M.H., and Sinha, S. (2013). Computational identification of diverse mechanisms underlying transcription factor-DNA occupancy. PLoS Genet. *9*, e1003571.

90. Ramos, A.I., and Barolo, S. (2013). Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. Phil. Trans. R. Soc. B *368*, 20130018.

91. Farley, E.K., Olson, K.M., Zhang, W., Rokhsar, D.S., and Levine, M.S. (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. Proc. Natl. Acad. Sci. USA *113*, 6508–6513.

92. Crocker, J., Preger-Ben Noon, E., and Stern, D.L. (2016). The Soft Touch: Low-Affinity Transcription Factor Binding Sites in Development and Evolution*117*, First Edition (Elsevier Inc.).

93. Kribelbauer, J.F., Rastogi, C., Bussemaker, H.J., and Mann, R.S. (2019). Low-affinity binding sites and the transcription factor specificity paradox in eukaryotes. Annu. Rev. Cell Dev. Biol. *35*, 357–379.

94. Paakinaho, V., Presman, D.M., Ball, D.A., Johnson, T.A., Schiltz, R.L., Levitt, P., Mazza, D., Morisaki, T., Karpova, T.S., and Hager, G.L. (2017). Single-molecule analysis of steroid receptor and cofactor action in living cells. Nat. Commun. *8*, 1–14.

95. Li, J., Dong, A., Saydaminova, K., Chang, H., Wang, G., Ochiai, H., Yamamoto, T., and Pertsinidis, A. (2019). Single-molecule nanoscopy elucidates RNA polymerase II transcription at single genes in live cells. Cell *178*, 1–3.

96. Donovan, B.T., Huynh, A., Ball, D.A., Patel, H.P., Poirier, M.G., Larson, D.R., Ferguson, M.L., and Lenstra, T.L. (2019). Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. EMBO J. *38*, 1–18.

97. Kurafeiski, J.D., Pinto, P., and Bornberg-Bauer, E. (2019). Evolutionary potential of cis-regulatory mutations to cause rapid changes in transcription factor binding. Genome Biol. Evol. *11*, 406–414.

98. Kulkarni, M.M., and Arnosti, D.N. (2003). Information display by transcriptional enhancers. Development *130*, 6569–6575.

99. Junion, G., Spivakov, M., Girardot, C., Braun, M., Gustafson, E.H., Birney, E., and Furlong, E.M. (2012). A transcription factor collective defines cardiac cell fate and reflects lineage history. Cell *148*, 473–486.

100. Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683–689.

101. Li, J., White, J.T., Saavedra, H., Wrabl, J.O., Motlagh, H.N., Liu, K., Sowers, J., Schroer, T.A., Thompson, E.B., and Hilser, V.J. (2017). Genetically tunable frustration controls allostery in an intrinsically disordered transcription factor. eLife *6*, 1–17.

102. Finzel, A., Grybowski, A., Strasen, J., Cristiano, E., and Loewer, A. (2016). Hyperactivation of ATM upon DNA-PKcs inhibition modulates p53 dynamics and cell fate in response to DNA damage. Mol. Biol. Cell. *27*, 2360–2367. https://doi.org/10.1091/mbc.E16-01-0032.

103. McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., et al. (2018). CellProfiler 3.0: next-generation image processing for biology. PLoS Biol. *16*, e2005970.

104. Schröder, S., Herker, E., Itzen, F., He, D., Thomas, S., Gilchrist, D.A., Kaehlcke, K., Cho, S., Pollard, K.S., Capra, J.A., et al. (2013). Acetylation of RNA polymerase II regulates growth-factor-induced gene transcription in mammalian cells. Mol. Cell *52*, 314–324.

105. Sharova, L.V., Sharov, A.A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M.S.H. (2009). Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. DNA Res. *16*, 45–58.

106. Chan, L.Y., Mugler, C.F., Heinrich, S., Vallotton, P., and Weis, K. (2018). Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. eLife *7*, 1–32.

107. Ferguson, H.A., Kugel, J.F., and Goodrich, J.A. (2001). Kinetic and mechanistic analysis of the RNA polymerase II transcription reaction at the human interleukin-2 promoter. J. Mol. Biol. *314*, 993–1006.

108. Kugel, J.F., and Goodrich, J.A. (1998). Promoter escape limits the rate of RNA polymerase II transcription and is enhanced by TFIIE, TFIIH, and ATP on negatively supercoiled DNA. Proc. Natl. Acad. Sci. USA *95*, 9232–9237.

109. Wissink, E.M., Vihervaara, A., Tippens, N.D., and Lis, J.T. (2019). Nascent RNA analyses: tracking transcription and its regulation. Nat. Rev. Genet. *20*, 705–723.

110. Mehta, G.D., Ball, D.A., Eriksson, P.R., Chereji, R.V., Clark, D.J., McNally, J.G., and Karpova, T.S. (2018). Single-molecule analysis reveals linked cycles of RSC chromatin remodeling and Ace1p transcription factor binding in yeast. Mol. Cell *72*. 875.e9–887.e9.

111. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies and stains		
Hoechst 33342	Thermo Fisher Scientific	Cat #62249
anti-HA 6E2 mAb	Cell Signaling	Cat# 2367; RRID:AB_10691311
anti-GFP D5.1 mAb (rabbit)	Cell Signaling	Cat# 2956; RRID:AB_1196615
anti-mouse, Alexa Fluor 488	Cell Signaling	Cat# 4408; RRID:AB_10694704
Prolong Gold Antifade	Thermo Fisher Scientific	Cat# P10144
Chemicals, peptides, and recombinant proteins		
Lipofectamine 3000	Thermo Fisher Scientific	Cat# L3000001
Polyethylenimine (PEI)	Polysciences	Cat# 23966-1
20% Paraformaldehyde	Fisher Scientific	Cat# 50-980-492
iTaq^{TM}Universal SYBR Green reagent	BioRad	Cat# 1725121
Protoscript II reverse transcriptase	New England Biolabs	M0368
oligo-dT primers	New England Biolabs	S1316
Puromycin Dihydrochloride	Gibco	Cat# A1113803
Critical commercial assays		
RNeasy Plus Mini Kit	Qiagen	Cat# 74134
QIAprep Spin Miniprep Kit	Qiagen	Cat# 27106
Deposited data		
Processed data used for the figures.	This study	https://github.com/rosamc/kinsyn-2021. https://doi.org/10.5281/zenodo.7015032
Experimental models: Cell lines		
HEK293FT	Thermo Fisher Scientific	Cat# R70007
Experimental models: Organisms/strains		
Human: 1X UAS 1X127 ZF BS minCMV eGFPd2 monoclonal #2	This study	MP796
Oligonucleotides		
Actin_fwd: GGCACCCAGCACAATGAAGATCAA	Finzel et al.[102]	N/A
Actin_rev: TAGAAGCATTTGCGGTGGACGATG	Finzel et al.[102]	N/A
eGFP_fwd: AAGTTCATCTGCACCACCG	This study	N/A
eGFP_rev: TCCCTTGAAGAAGATGGTGCG	This study	N/A
zf_fwd: TTTTCGAGAAGACACGGCCT	This study	N/A
zf_rev: GCTGCTGTGGTCGGAGAAAT	This study	N/A
Recombinant DNA		
1X UAS 1X127 ZF BS minCMV eGFPd2	This study	pKC75
pUBC ZF127(WT)-BRD4	This study	pKC9
pUBC ZF127(WT)-cMyc	This study	pKC10
pUBC ZF127(WT)-HSF1	This study	pKC11
pUBC ZF127(WT)-HSF1mut1	This study	pKC12
pUBC ZF127(WT)-SP1	This study	pKC14
gRNA_AAVS1-T2 plasmid	Addgene	41820
VP12 humanSpCas9-Hf1 plasmid	Addgene	72247
Software and algorithms		
Python 3 (3.6, 3.8)	Python Software Foundation	https://www.python.org/
FlowJo 10.6.2	FlowJo LLC	https://www.flowjo.com

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cell profiler v 4.2.0	McQuin et al.[103]	https://cellprofiler.org
Scripts for running simulations and analyzing results	This paper	https://github.com/rosamc/kinsyn-2021. https://doi.org/10.5281/zenodo.7015032
Code and scripts to write the c++ code to calculate the steady-state of the model, and run the exploration of the boundaries of a regulatory strategy.	This paper	https://github.com/rosamc/GeneRegulatoryFunctions https://doi.org/10.5281/zenodo.7015068
Accessory code.	This paper	https://github.com/rosamc/scriptssyn https://doi.org/10.5281/zenodo.7011467
Other		
Goat serum	VWR	N/A
Attune NxT Flow Cytometer	Thermo Fisher Scientific	Attune NxT

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Angela DePace (Angela_DePace@hms.harvard.edu).

Materials availability

All DNA constructs and cell lines are available from A.S.K. and A.H.D.

Data and code availability

- Processed data is available at https://github.com/rosamc/kinsyn-2021 (the DOI is listed in the key resources table). The raw image data and all other data reported in this paper will be shared by the lead contact upon request.
- The boundary search code is available at https://github.com/rosamc/GeneRegulatoryFunctions. The rest of the code to reproduce the calculations and figures in the paper is available at https://github.com/rosamc/kinsyn-2021 and accessory code is available at https://github.com/rosamc/scriptssyn. The corresponding DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture

HEK293FT cells (Thermo Fisher Scientific) were used as a background cell line in this study. Cells were cultured in DMEM with L-glutamine, 4.5g/L Glucose and Sodium Pyruvate (Thermo Fisher Scientific) supplemented with 10% FBS (Clontech), GlutaMAX supplement (Thermo Fisher Scientific), MEM Non-Essential Amino Acids solution (Thermo Fisher Scientific) and 1% penicillin-streptomycin (Thermo Fisher Scientific). Cells were maintained at 37°C with 5% CO2 in a humidified incubator, with splitting every 2-3 days.

METHOD DETAILS

Modelling details and the linear framework

In this work we have used the linear framework formalism to model the interplay between the binding of TFs and their effects on the transcription cycle. This framework was introduced in Gunawardena et al.,[59] and we have previously exploited it to study other problems in gene regulation.[27,31,60–62] can be consulted for details. We outline the main features here.

A biological system is represented by a finite, directed, labelled graph G with labelled edges and no self-loops. The graph represents a coarse-grained version of the system of interest, with the nodes being the states of interest, and the edges the transitions between them. The edge labels are the infinitesimal transition rates for the underlying Markov process, with dimensions of $(\text{time})^{-1}$, and they include terms that specify the interactions between the graph and the surrounding environment. For example, the transitions that represent the binding of a TF have edge labels that include the TF concentration, which is assumed to remain constant over time (i.e., TF is sufficiently in excess that, to a good approximation, binding does not reduce the concentration of free TF available for binding).

The graph defines the time-evolution of the probability for each state of the system (vertex) as follows. Assume that each edge is a chemical reaction that follows mass-action kinetics with the edge label as the rate. Since each edge has only one source vertex, the resulting dynamics is linear and is described by a matrix equation,

$$\frac{d\vec{P}}{dt} = \mathcal{L}(G)\vec{P}. \quad \text{(Equation 4)}$$

Here, $\vec{P}$ is the column vector of state probabilities at time t, with dimension n, and $\mathcal{L}(G)$ is the Laplacian matrix of the graph. Equation 4 is the master equation, or Kolmogorov forward equation, of the underlying Markov process.

For a strongly connected graph, the system has a unique steady state, where $d\vec{P}/dt = 0$. The steady-state probability values for each state are computed by summing over the products of the rate labels for each of the spanning trees rooted at that state, and normalising appropriately (see Estrada et al.,[27] for details).

The mRNA concentration m is assumed to evolve according to:

$$\frac{dm}{dt} = k_{3,\varnothing}P_{3,\varnothing} + k_{3,A}P_{3,A} + k_{3,B}P_{3,B} - \delta_m m \quad \text{(Equation 5)}$$

where the $P_{3,X}$ are the probabilities of states $3_\varnothing$, 3_A, 3_B at a given time (Figure 1B). By assuming steady state, setting $dm/dt = 0$, and dividing by δ_m, we obtain the expression for the steady-state mRNA (* denotes steady state):

$$m^* = \frac{k_{3,\varnothing}}{\delta_m}P^*_{3,\varnothing} + \frac{k_{3,A}}{\delta_m}P^*_{3,A} + \frac{k_{3,B}}{\delta_m}P^*_{3,B} \quad \text{(Equation 6)}$$

$$m^* = \overline{k_{3,\varnothing}}P^*_{3,\varnothing} + \overline{k_{3,A}}P^*_{3,A} + \overline{k_{3,B}}P^*_{3,B} \quad \text{(Equation 7)}$$

This gives Equation 1 of the main text, where the overbars are dropped for simplicity. In the parameter exploration, we directly sample the normalised rates.

Biologically plausible ranges for parameters

We considered a biologically plausible range for the normalised parameter values to be between 1 and 10^4, according to the following reasoning:

The events from the binding of the polymerase complex until the production of an mRNA molecule involve many biochemical reactions, including the binding interactions associated with the assembly of the pre-initiation complex, the phosphorylation of the C-terminal domain of RNA polymerase and other post-translational modifications,[104] as well as the biochemistry associated to elongation. Our 3-state cycle is therefore a coarse-grained representation of all these processes. In order to determine biologically plausible parameter ranges, we searched for measurements of reaction rates for these processes, and normalised those to typical rates of mRNA degradation, taken to have typical half-lives between 10 min (0.00116 s^{-1}) and 5 h (3.85×10^{-5} s^{-1}).[105,106]

For a reaction at a rate of 0.7 s^{-1} (~1 s half-life), normalizing by the mRNA degradation rates would result into a normalised range of 600–18000.

For a rate of 0.07 s^{-1} (~10 s half-life), the normalised range would be 60-1800.

For a rate of 0.016 s^{-1} (~1 min half-life), the normalised range would be 10-300.

And for a rate of 0.00116 s^{-1} (~10 min half-life) the normalised range would be 1-30.

These values are consistent with measurements of various transcription-associated biochemical reactions: the *in vitro* rate of pre-initiation complex assembly was found to vary over ranges on the order 10^{-3} s^{-1}[107] to 0.1 s^{-1},[108] and the rate of promoter opening/escape was reported to be 0.002 s^{-1}.[108] Pause stability is estimated to be from 3 s to 20 min.[109] And the TF residence time can be from just a few seconds to a few minutes.[94,110]

Therefore, we took a range of 1-10^4 for our parameter values. We note that for smaller ranges representing slower rates for the basal polymerase cycle and constraining the fold-change by which a given TF can modify a rate, the results are qualitatively the same, but the synergy space region is reduced (Figure 2B, gray).

Synergy boundary for a regulatory strategy

In order to determine the region of the synergy space that can be spanned by a given regulatory strategy, we used a biased random sampling algorithm, modified from that in Estrada et al.[27] Briefly, this algorithm iteratively samples parameter sets with the goal to fill as much of the synergy space region as possible. Parameters were chosen from a given range of normalised rate values, and TFs were assumed to at most modify the basal rates by a certain factor (see figure captions for the values corresponding to each figure). A maximum fold change for expression in the presence of one TF alone (at 2× concentration) was also pre-specified, such that parameter sets that generate expression outside this range were discarded. The steps of the algorithm are the following:

1. The algorithm is initialised with the corresponding hyperparameters (below) and the constraints that will need to be kept among parameter values during sampling (e.g. fold change by which a given TF can modify a given basal rate). The synergy space

region is divided into a grid, so that each grid cell corresponds to a small region of the synergy space, which we consider a point.

2. A few parameter sets are sampled until 10 grid cells are occupied. Each parameter set is sampled in two steps. First, the parameters for the basal transitions in the cycle and the TF binding and unbinding are sampled uniformly on a logarithmic scale from their specified range (e.g. if the parameter range is $[10^a, 10^b]$, we choose a number v uniformly from the interval $[a, b)$ and then take 10^v as the parameter value). Then the constraint parameters are sampled similarly, but with their ranges adjusted according to the chosen value for the reference parameter. The current working boundary is defined as those grid cells with only empty cells above or below in the same column, or to the right and left on the same row.
3. Iteratively until convergence: each boundary cell provides a parameter set that is slightly modified in different ways to search its neighborhood, as explained below. New parameter sets that generate points outside the current working boundary are always kept for the subsequent iteration, and those that generate points at the current working boundary are kept with some probability (*prob_replace*). The current working boundary is recomputed after visiting all parameter sets in the previous working boundary. Convergence is determined by 3000 iterations without changes in the working boundary.

In order to search the surrounding parameter space of a given parameter set (point in synergy space), we followed 3 strategies. At each iteration, each strategy was applied to all points before moving to the next one, provided there were sufficient points for strategies 2 (10) and 3 (100):

1. Randomly select a few parameter values and modify them (“mutate” them).
2. “Pull” towards a target point in the direction determined by the centroid and the point being modified, away from the boundary: for at most 500 trials, slightly modify the parameter set, and keep the new one if it generates a point in synergy space closer to the target. Stop when the distance to the target is smaller than a tolerance.
3. “Pull” in the direction (approximately) perpendicular to the tangent between the point being modified and its neighbor, as in 2.

The two pulling strategies help to intensely explore the parameter space around a given parameter set and allow the boundary to escape from local minima regions.

The algorithm depends on various hyperparameters: probability of selecting a parameter value for mutation (we used 0.2 and 0.5), probability of replacing an already-existing boundary parameter set if a new one generates a synergy value that falls in the same grid cell (we used 0.2 and 0.6), width of the interval around a parameter value to sample for new parameter values (we used, in base 10 logarithmic scale: [-2,2],[-1.5,1.5],[-1,1]). Searches were run for all 12 combinations of hyperparameters, and results were merged together.

Random sample of points in synergy space

In order to randomly sample parameter values in the synergy space (Figure S1C) we followed a rejection sampling approach. Parameters were sampled logarithmically from their predefined range (1-10^4) (i.e. for each parameter value w, we obtained a parameter value v uniformly from the interval [0,4), and set $w = 10^v$.) The constraints on the maximum fold change effect on the polymerase cycle rates by the TFs were checked, as well as the constraint on the expression fold change by each of the TFs at 2×. Only parameter sets that satisfied all constraints were kept. Then, in order to have a more uniform distribution of points over the synergy space, we binned the synergy space into a grid with bins every 0.025 $S_{A,B}$ and $S_{B,A}$ units, and kept one parameter set per bin. For Figure S1Ci, we collected 1 million parameter sets that satisfied the constraints, kept one parameter set per bin (the first one encountered that would correspond to a given bin), and repeated the procedure 10 times. For Figures S1Cii and S1Ciii, we just evaluated 10 million parameter sets and kept one per bin.

Effect of binding/unbinding rate on synergy

In order to explore how synergy depends on the binding and unbinding rates of the TFs (Figures 4 and S6), we generated sets of basal parameters by randomly sampling on a logarithmic scale the basal rates between 1 and 10^4, and the binding and unbinding rates between $10^{1.5}$ and 10^3. For each of these basal parameter sets, we generated 1000 parameter sets corresponding to the TF-associated parameters that satisfy some constraints regarding the strength of the TFs. For this, we obtained parameter values for the TF-associated parameters by random sampling on a logarithmic scale, also between 1 and 10^4, and then checked whether or not the following constraints were satisfied: i) TF-associated parameter values at most a thousand-fold higher than the respective basal ones (or a thousand-fold lower for counterclockwise rate k_4); ii) fold change in expression by each TF individually at 2X concentration between 1 and 5; iii) TF A is consistently the strongest of the pair when the binding or unbinding rate is changed by a factor f, where f spans 10 logarithmically spaced values over two orders of magnitude, between 0.1 and 10 ($f = 10^{-1+0.22i}$, $i = 0, 1, 2 \cdots 9$). For each basal parameter set, we kept 1000 TF-associated parameter sets that satisfy these constraints, corresponding to 1000 complete parameter sets. For each of these parameter sets, we determined the class of behaviour in synergy space as a function of the change in the binding or unbinding rate over the above-described two-order magnitude range as described in the main text, and saved for downstream analysis those parameter sets where the absolute value of the change in both $S_{A,B}$ and $S_{B,A}$ was at least 0.05.

Construct design and cloning

The reporter construct consists of a single synthetic zinc finger binding site (CGGCGTAGCCGATGTCGCGC) upstream of a minimal CMV promoter (taggcgtgtacggtgggaggcctatataagcagagctcgtttagtgaaccgtcagatcgcctgga) driving d2EGFP (EGFP destabilized with signal peptide for fast degradation (fusion with aa 422-461 of mouse ornithine decarboxylase)).

synTF fusion proteins containing an activation domain of interest fused to an N-terminal zinc-finger binding domain with a GGGGS flexible linker were driven under control of a ubiquitin promoter and contain a 5' sv40 nuclear localization sequence, C-terminal HA and rabbit globin polyA 3' UTR. Genome-orthogonal zinc fingers were previously developed to target 20-bp sequences that minimize identity with the reference human genome.[57,58] To tune the affinity of the zinc finger, 5, or 7 arginine residues in the zinc finger array were mutated to alanine as denoted in Figure S7. The following protein domains were selected and conjugated as respective activation domains according to previous studies:

SP1 (Residues 263 – 499) [PMID: 8278363]
NITLLPVNSVSAATLTPSSQAVTISSSGSQESGSQPVTSGTTISSASLVSSQASSSSFFTNANSY
STTTTTSNMGIMNFTTSGSSGTNSQGQTPQRVSGLQGSDALNIQQNQTSGGSLQAGQQKE
GEQNQQTQQQQILIQPQLVQGGQALQALQAAPLSGQTFTTQAISQETLQNLQLQAVPNSGP
IIIRTPTVGPNGQVSWQTLQLQNLQVQNPQAQTITLAPMQGVSLGQTSSSN
cMyc (Residues 1-70) [PMID: 12177005]
MDFFRVVENQQPPATMPLNVSFTNRNYDLDYDSVQPYFY
CDEEENFYQQQQQSELQPPAPSEDIWKKFEL
BRD4 (Residues 1308-1362) [PMID: 24860166]
PQAQSSQPQSMLDQQRELARKREQERRRREAMAATIDMNFQSDLLSIFEENLF
HSF1 (Residues 370-529) [PMID: 9606196]
PEKCLSVACLDKNELSDHLDAMDSNLDNLQTMLSSHGFSVDTSALLDLFSPSVTVPDMSLP
DLDSSLASIQELLSPQEPPRPPEAENSSPDSGKQLVHYTAQPLFLLDPGSVDTGSNDLPVLF
ELGEGSYFSEGDGFAEDPTISLLTGSEPPKAKDPTVS
HSF1 mutant (Residues 370-529, F418A, F492A, F500A) [PMID: 9606196]
PEKCLSVACLDKNELSDHLDAMDSNLDNLQTMLSSHGFSVDTSALLDLASPSVTVPDMS
LPDLDSSLASIQELLSPQEPPRPPEAENSSPDSGKQLVHYTAQPLFLLDPGSVDTGSNDLP
VLAELGEGSYASEGDGFAEDPTISLLTGSEPPKAKDPTVS

Genomic integration of reporter constructs

Reporter lines were generated by site-specific integration of reporter constructs into HEK293FT cells using CRISPR/Cas9 mediated homologous recombination into the AAVS1 (PPP1R2C) locus as previously described.[57] Briefly, 60,000 cells were plated in a 48-well plate and transfected the following day by PEI with a mixture of the following: 70ng of gRNA_AAVS1-T2 plasmid (Addgene 41820), 70 ng of VP12 humanSpCas9-Hf1 plasmid (Addgene 72247), and 175 ng of donor reporter plasmid. Donor reporter plasmids contain flanking arms homologous to the AAVS1 locus, a puromycin resistance cassette, and constitutive mCherry expression. After transfection, cells were cultured in 2 mg/mL puromycin selection for at least 2 weeks with splitting 1:10 every 3 days. Monoclonal populations for reporter cell lines were isolated by sorting single cells from this population into a 96-well plate and growing cell lines from each well. A minimum of 6 monoclonal cell lines that express high level of mCherry protein were transiently transfected with a strong synTF activator (HSF1 or VP16) and a monoclonal cell line to be used going forward was selected based on the fold-change of GFP expression relative to basal GFP level.

Transient transfection

Stable reporter cell lines were transfected with synTF plasmid constructs using polyethylenimine (PEI, Polysciences) as described in Park et al.[57] 60,000-100,000 cells/well were plated in 48-well plates and transfected the following day with a total of 10ng per synTF, unless otherwise noted, with single stranded filler DNA (Thermo Fisher Scientific) up to 200ng total. 50ng of pCAG-iRFP720 (Addgene, #89687) was used as a transfection control plasmid in flow cytometry experiments. Two days after transfection, cells were collected and prepared for flow cytometry, unless otherwise noted.

Flow cytometry and data analysis

For each measurement, cells were harvested and run on an Attune NxT (Thermo Fisher Scientific) Flow Cytometer equipped with a high-throughput auto-sampler. A minimum of 10,000 events were collected for each well and were gated by forward and side scatter for live cells and single cells, as described in Park et al.[57] Cells were then gated by iRFP for transfection-positive populations. The geometric mean of GFP fluorescence distribution was calculated in FlowJo (Treestar Software). GFP expression fold-change was determined by normalizing with mean GFP intensity of the reporter only control. Flow cytometer laser/filter configurations used in this study were: EGFP (488 nm, 510/10), mCherry (561 nm, 615/25), iRFP-720 (638 nm, 720/30). All flow cytometry measurements were performed in technical replicates. Considering together all replicates from all experiments with the same transfection condition, we checked for consistency and discarded technical errors. This removed the cMyc 2X condition in one of the experiments since it

yielded an aberrantly low fold change. Moreover, we removed 4 additional replicates, each from a different condition, that had a fold change that was above/below two standard deviations from the mean considering all replicates for that condition together.

Western blotting

A reporter cell line was transfected with indicated amounts of ZF-HSF1 (0, 10, 20, 50, 100, 150, 200ng) in a 48-well plate at a cell density of 1x105 per well. After 2 days, cells were rinsed with PBS and lysed with 200 μL of NuPAGE LDS sample buffer (Thermo Fisher Scientific), followed by 5 seconds of sonication. Whole cell lysates were mixed with NuPAGE Sample Reducing agent (10X, Thermo Fisher Scientific) at 95°C for 5 minutes. Samples were then loaded into a 4-12% NuPAGE Bis-Tris Mini Protein precast gel (Thermo Fisher Scientific) and were run at 200V for 30 minutes in NuPAGE MES SDS Running Buffer. Separated proteins were transferred to a PVDF membrane using P0 protocol of iBlot2 system (Thermo Fisher Scientific). Membranes were blocked for 1hr at room temperature in blocking solution (5% w/v nonfat dry milk in 1X PBST) with gentle rocking. The membranes were probed with anti-HA (1:4000; Abcam ab9110) and anti-GAPDH (1:1000; Abcam ab9485) antibodies at room temperature for 1 hour with gentle rocking. The membranes were washed in PBST three times for 5 minutes each, and incubated with a goat anti-rabbit IgG-HRP antibody (1:2000; Abcam ab6721). The target proteins were visualized by chemiluminescence using SuperSignal West Pico PLUS substrate (Thermo Fisher Scientific) and an iBright Western Blot Imaging Systems (Thermo Fisher Scientific). Quantification of band intensities was carried out using FIJI.[111]

QUANTITATIVE REAL-TIME PCR

1×10^5 Hek293FT reporter cells were seeded one day prior to transfection in 6cm culture dishes. Transfection was performed with the indicated amounts of synTF plasmid as described above for flow cytometry experiments using polyethylenimine (PEI) (polyscience) or Lipofectamine 3000 (Thermo Fisher Scientific). Two days post transfection, cell pellets were harvested and mRNA was extracted using the RNeasy Mini Kit (Qiagen). 500 ng extracted total RNA was reverse transcribed into cDNA for each sample. Reverse transcription was performed using Protoscript II reverse transcriptase (New England Biolabs) and oligo-dT primers (New England Biolabs). Quantitative real-time PCR was performed in triplicates using iTaq™Universal SYBR®Green reagent (Bio-Rad) on a CFX96 PCR machine (Bio-Rad). Primers were used in a final concentration of 243.2 nM. β-actin expression was used as a reference gene for relative quantification of RNA levels. Used primer sequences are (5'-3'):

Actin_fwd: GGCACCCAGCACAATGAAGATCAA;
Actin_rev: TAGAAGCATTTGCGGTGGACGATG;
eGFP_fwd: AAGTTCATCTGCACCACCG;
eGFP_rev: TCCCTTGAAGAAGATGGTGCG;

Immunofluorescence to assess synergy for cMyc-SP1

Immunostaining

0.5×10^5 cells were seeded on poly-Lysine coated high-precision glass coverslips (18 mm round, #1.5) in 12-well culture plates one day prior to transfection. Transfection was performed as described for flow cytometry experiments. A total amount of 200 ng DNA (20 ng synTFs and 180 ng ssDNA) was used for transfection experiments. PEI was scaled to 12-well plate volume of 100 μL total transfection mix. 48 h post transfection, cells were washed with 1x PBS, fixed with 2% PFA (Fisher Scientific) and blocked for 30 min with 10% Goat serum (VWR) in 1x PBS after washing. Immunodetection was performed using anti-HA-tag (6E2) mouse monoclonal antibody (Cell Signaling) and anti-GFP (D5.1) rabbit monoclonal antibody (Cell Signaling) 1:200 in 1%BSA/PBS overnight. Cells were washed with 0.1% Triton X-100 and incubated with anti-mouse IgG Alexa Fluor 488 (#4408, Cell Signaling) and anti-rabbit IgG Alexa Fluor 647 antibodies 1:1000 in 1%BSA/PBS for 1 h. After washing with 0.1% Triton X-100, nuclei were stained with 2 μg/mL Hoechst-33342 (Thermo Fisher Scientific) and mounted on glass slides using Prolong Gold Antifade (Thermo Fisher Scientific). Image acquisition was performed at least 16 h after mounting coverslips on glass slides.

Fluorescence microscopy

Images were acquired as single-plane multipoint positions on a Nikon Ti2 inverted microscope upon illumination by a Lumencor Sola 395 Light Engine and a Plan Apo VC 20x objective (NA 0.75). The following filter sets were used. Alexa Fluor 488: excitation FF01-466/40, emission FF03-525/50, dichroic FF495-Di03 (all Semrock); Hoechst-33342: excitation ET395/25x, emission ET460/50m, dichroic ET425lp (all Chroma). Detection was performed with a Hamamatsu ORCA Flash 4.0 LT camera. NIS elements software for image acquisition was used.

Image Processing

Images were extracted from nd2 files, separated as.tif-files per channel and field of view. CellProfiler 4.0[103] was used for image segmentation and measuring nuclear fluorescence intensity. A pipeline was customized based on the pipeline for Human cells provided by the CellProfiler project. Nuclei segmentation was performed based on Hoechst-33342 staining using adaptive Otsu thresholding and a nuclear diameter range of 30 – 50 pixels. Objects at the border of images and outside this range were excluded. Clumped nuclei were distinguished based on fluorescence intensity and object intensity was calculated for the segmented nuclear area in all channels. For simplicity, GFP fluorescence in the nuclear area was used as a proxy for overall GFP expression in cells. The percentage of

properly identified nuclei is different between images due to segmentation challenges from dense and overlaying cell growth of Hek293T cells that could not be fully resolved by adjusting parameters. From each results table, the integrated intensity of all three channels was used for further analysis (Intensity_IntegratedIntensity_DNA, Intensity_IntegratedIntensity_GFP, Intensity_IntegratedIntensity_TF).

Data analysis

We began by removing outliers based on the GFP signal per cell, which can represent segmentation errors or unhealthy cells. For each transfection condition and the control condition (untransfected), we removed the cells with the 5% lowest and 5% highest GFP signal. For the control condition, the HA-tag signal fluorescence corresponding to the remainder cell population was used to define the cutoff for positive transfection. In Figure S5, we show data where we selected the 90th percentile of the control signal as the cutoff, but very similar results were obtained with a more stringent cutoff corresponding to the 95th percentile.

Given the variability associated with transfection, the different transfection conditions result in similar distributions for the HA-tag signal which differ in their tails, with the cMyc condition spanning the lowest range in the input levels (Figure S5A). As a result, the input in the combination is statistically significantly higher than in the cMyc condition at a significance level of 0.005 according to two non-parameteric tests that are applicable in this situation: the one-sided Kolmogorov-Smirnov test and the one-sided Mann-Whitney U test (as implemented by the Python Scipy library, version 1.6.2). In order to ensure that we calculate synergy for cells where the input levels in the combination are not higher than in each of the individual transfected conditions (in particular, the cMyc condition), we selected the highest input fluorescence level for which significance is lost for both tests (considering a 0.005 significance level). This cutoff encompasses 91% of the data for the cMyc condition. We then calculate synergy for the subpopulation of cells with input levels at or below this cutoff (yellow region in Figure S5A).