Title

Revisiting Mechanisms of Eukaryotic Gene Regulation

Author

Karr, Jonathan Paul

Publication Date

2023

Peer reviewed|Thesis/dissertation

Revisiting Mechanisms of Eukaryotic Gene Regulation

By

Jonathan Paul Karr

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cellular Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:
Professor Xavier Darzacq, Co-Chair
Professor Robert Tjian, Co-Chair
Professor Dirk Hockemeyer
Professor Elcin Unal
Professor Gary Karpen

Fall 2023

Abstract

Revisiting Mechanisms of Eukaryotic Gene Regulation

by

Jonathan Paul Karr

Doctor of Philosophy in Molecular and Cellular Biology

University of California, Berkeley

Professors Xavier Darzacq and Robert Tjian, Co-chairs

Cis–trans relationships govern gene regulation: it is the interactions between diffusible proteins and chromatinized DNA that dictate nuclear functions. The nature of those interactions must allow for the great regulatory complexity needed for eukaryotic function and particularly of multicellular development. However, many of the current models of cis–trans relationships in the nucleus either lack biophysical rigor or have gone untested. This thesis of three chapters addresses three shortfalls in the current conceptualization of how gene-regulatory proteins function.

In the introductory chapter, it is argued that there is a widespread over-reliance on qualitative descriptions of biomolecular behaviors and functions. Because biochemistry is inherently stochastic, the way in which such behaviors and functions are described must eventually be probabilistic and quantitative. Simplistic 1:1 relationships between molecules need to be dropped in favor of probabilistically constituted ensembles of factors with different affinities. Furthermore, because the cell interior is literally fluid and molecular interactions quite transient, the temporal dynamics of the system must always be kept in mind, which largely forbid the use of static structures and monolithic series of events.

The second chapter issues a challenge to the notion that an important cis relationship, that between enhancers and promoter, is mediated by a direct interaction between the two elements—a striking instance of the structure–function paradigm questioned in the first chapter. It has been assumed for decades that a complex of proteins forms a bridge connecting enhancer to promoter even over large genomic distances. However, neither evidence nor reason rules this assumption in, and other models have hardly been considered, much less ruled out. Here it is proposed that diffusible biochemical species in the form of modified proteins are generated at the enhancer and can affect the promoter via diffusion. What determines the scale of the gradient of signal emanating from the enhancer is the balance of the local generation rate and the ubiquitous degradation rate.

Finally, an imaging study is presented in which the question is asked, how are transcription coactivators (trans factors) distributed among the plethora of cis elements? Through single-molecule tracking of p300 and many mutants thereof, it is shown that direct cis interaction—i.e., binding modified histones—is not responsible for coactivator targeting. Rather, p300 depends on a combination of transcription factor–interaction domains to associate with chromatin, indicating that atop the regulatory layer of cis–trans interactions between transcription factors and chromatin is another in the form of trans interactions between transcription factors and their cofactors.

Table of Contents

Chapter 1: Reconsidering Models of Biomolecular Interactions

Macroscopic analogies are at once indispensable and inimical to understanding things at the molecular scale—indispensable because our language is not so abstract as to be able to jettison them entirely, yet inimical because the intuitions they engender are usually wrong. It requires deliberate mental effort to disabuse oneself of such false intuitions. This essay along with the thesis which it introduces represents a nascent (and therefore stumbling) attempt at understanding nuclear function with fewer macroscopic biases than are in evidence throughout the present literature.

The prevailing framework for understanding eukaryotic gene regulation is rife with intuitions derived from common experience in the macroscopic world. Salient linguistic examples include speaking of molecules as components of "machines", as being "recruited" by other molecules, as having "structure", as "assembling", as having an assignable "function", as carrying out "programs". Molecules are erroneously conceptualized as purposive players moving from place to place to perform their role in the greater narrative (see "Is the cell really a machine?"[1]).

It must be acknowledged at the outset that such informal language belies a great deal of sophistication ready to the molecular biologist's hand. It is well known, for instance, that molecular motion is random and diffusive, not purposive or directional. It is quite understood that the interactions of biomolecules are governed by statistical mechanics, that they are probabilistic and not deterministic, that they are dynamic instead of static, and so on. It is for the sake of simplicity and intuitiveness that much informal language has been adopted. But indulging in such linguistic conveniences has come at a cost.

Consider the current understanding of the sequence of events which results in a gene being turned on. (Consider first that it is entirely permissible to speak of a singular sequence of events in this context and of a gene being "turned on.") It is understood that a transcription factor—a molecule whose "function" it is to turn certain genes "on" or "off"—"searches" until it "finds" a "target" promoter by "recognizing" a specific DNA sequence where it will "bind" and "recruit" other regulatory factors until eventually the transcriptional "machinery" is "assembled" to initiate transcription (quotes here denote not sarcasm but deliberate usage of macroscopic phraseology common in the academic literature). Elaboration of this template narrative with context-specific details constitutes the "mechanism" of a particular gene-regulatory pathway.

All the terms in the previous paragraph bracketed by quotation marks arguably run afoul of principled molecular-scale thinking, or at least have the potential to; yet they are used ubiquitously in formal publications to interpret data and are even translated into pictorial models representing up-to-date conceptualizations of molecular processes. It is therefore difficult to maintain that they are mere conveniences and not functional concepts in the minds of active researchers. Hopefully, conscious recognition of one's dependency on such language will provide new and less problematic conceptual avenues. It is with such hope in view that the following attempt is made at articulating some familiar processes with somewhat less familiar words.

A useful starting point for this project is the recently published work from the Hager group showing that transcription factor (TF) dwell times on chromatin are power-law distributed, implying a continuum of affinities rather than a set of discrete affinities.[2] As the authors argue forcefully, this finding negates the widespread notion that there are two discrete categories of TF–DNA binding: specific and non-specific. Instead, it suggests that while a TF is diffusing in the nucleus, it will be temporarily detained to various degrees by *all* DNA sequences. That is, its rate of diffusion will be

non-uniform depending on the strengths of interactions in different chromatin regions. It is not "recruited" to any DNA sequence; it binds all DNA sequences with continuously varying affinities. It is not "searching" for "targets"; when it collides with a DNA sequence, it is temporarily detained for a length of time that is related probabilistically to its affinity for that sequence. With respect to a given TF, the chromatin is a diffusion-modifying landscape, with certain DNA sequences seemingly evolved to detain it for longer—that is, to achieve a higher relative time-averaged occupancy—such as at promoters of genes whose transcription depends on the TF. However, the TF and its associated biochemical activities are not limited to such sites, which differ only quantitatively from what are regarded as non-specific or off-target binding sites. Nor is evolution concerned with human classifications of correct and incorrect binding events; it can—and indeed must—operate on the whole probabilistic landscape of TF–DNA interactions.

According to the canonical understanding, once a TF "finds its target," it functions to "recruit" a cadre of additional factors via specific protein–protein interactions (PPIs) ultimately to "assemble" the transcriptional "machinery". In reality, those PPIs will behave much like the protein–DNA interactions that determine the distributions of TFs in the nucleus: they do not consist of exquisitely specific, one-to-one biomolecular recognition events resulting in static recruitment; rather, the PPIs will detain the randomly diffusing coactivators for shorter or longer intervals (microseconds to seconds) depending on their mutual affinities. And the so-called transcriptional machinery is not a stereotyped and stoichiometric mechanical structure but a probabilistic ensemble of various and partially redundant molecules which likely do not even have to be simultaneously present.[3–5]

Embracing rather than ignoring the stochasticity and intrinsic promiscuousness of biomolecular interactions yields a conceptualization of the nucleus as a swarm of diffusing molecules that are being detained for different lengths of time in a spatially biased manner determined by a sequenced and heritable substrate—DNA. The proteins that interact with DNA sequences with differential affinities—the TFs—will secondarily bias their PPI partners to also be detained for varying intervals at different DNA sites in a probabilistic manner. Consequently, even at steady state, when the nucleoplasm is well mixed, there is not a homogenous mixture of molecules; rather, there is a spatially biased probabilistic distribution of DNA-binding TFs and their interactors throughout the nucleus. (Notice how difficult it is to draw a picture of this conceptualization or to assign discrete relationships between different molecular "players.")

Under the prevalent molecular biology paradigm, when such a spatial pattern is observed, it is assumed to arise from structure—that is, from a molecular assembly whose constituents have evolved to ensure its reproducible and stable configuration. That cognitive bias is arguably the underlying reason for the explosive enthusiasm throughout the past decade over phase-separating phenomena: they provided what is essentially a structural explanation for inhomogeneous distributions of molecules in liquid phases in the form of a meshwork of multivalent interactions. Because TFs often exhibit such distributions in vivo, and because there were in vitro demonstrations that spatial inhomogeneities could arise spontaneously through phase separation from homogeneous mixtures of TFs and DNA, it became popular to ascribe the spatial patterns observed in the nucleus to large structures held together by networks of multivalent interactions.[6] Likewise, when it was observed that accessible regions of chromatin had inhomogeneous nuclear distributions, it was assumed that such arose from protein-scaffolded structures.[7]

The more parsimonious and flexible alternative to such structural models is that inhomogeneous distributions of nuclear factors are patterns emerging from positionally retarded diffusion of TFs

due to their interactions with chromatin. If there is something in that which could be called a structure, it is not at all like a macroscopic building with its fixed components in rigid relationships; it is a probabilistic ensemble of dynamically exchanging biomolecules. Those molecules have evolved not as components of one structure but as constituents of various ensembles that will arise in diverse contexts, which could be why nearly all of them have intrinsically disordered regions—peptides capable of flexible (diverse) interactions.[8]

However, it is not the mere distribution of molecules or their passive binding interactions that hold interest in understanding the chemical system that is the nucleus; it is the chemical reactions that are taking place and how the distributions of enzymes and substrates affect catalysis. Indeed, the spatial heterogeneity of the nucleus is suggestive of the potential to perform *patterned* and *patterning* reactions. Consider Turing's remarkable demonstration that a mere discrepancy in diffusion rates of an activator and an inhibitor can suffice for forming stable patterns even from an initially homogenous system.[9] Consider also Wolpert's later realization that in a biological system with inhomogeneous starting conditions (such as the DNA sequence in the nucleus, which has rich variation across its vast length), the concentration of a morphogen could be converted into positional information—i.e., spatial patterns—via sensory thresholds.[10] By extension, could DNA nucleate regulatorily meaningful spatial patterns in the nucleus via biased diffusion of TFs, localized activating and inhibiting reactions, and some manner of threshold-sensing?

The immediate barrier to that proposition is the fact that in both Turing's and Wolpert's models the morphogens were being produced and degraded. Although protein metabolism and diffusion occur on a spatiotemporal scale relevant for morphological development, proteins are too long-lived and fast-diffusing to create gradients within a nucleus in a similar manner. Post-translational modifications (PTMs) could serve to fill this spatiotemporal gap (Chapter 2). Because PTM depositions, such as acetylation and phosphorylation, are widely catalyzed in the nucleus by TF-interacting cofactors, whereas PTM removal is catalyzed by enzymes both ubiquitous and abundant, there is the theoretical potential for gradients to form in the nucleus, albeit at a scale difficult to probe experimentally. Hence, it seems there is a spatial–chemical regime in the nucleus that would enable reaction-diffusion mechanisms.

Perhaps some would view the proposition of reaction–diffusion being a prevalent phenomenon in the nucleus with skepticism, but upon careful inspection the presumption of its absence or insignificance appears the more dubious alternative. To hold that reaction-diffusion is *not* at play requires embracing the implausible notion that the products of all reactions that occur at cis-regulatory elements (CREs) are functionally confined to their immediate locale (within 10s of angstroms) because they are unable either to diffuse elsewhere or to exert any effect if they do. It could be that the longtime but mistaken focus on histones as the substrate of PTM depositors has disposed the field to thinking that reactions in the nucleus are so spatially confined. Little else would account for it. Once it is acknowledged that diffusible substrates—most notably TFs—are modified by so-called epigenetic writers (e.g. "histone" acetyltransferases, methyltransferases), the reaction-templating potential of chromatin is immediately appreciable.

But what have all these considerations to do with gene regulation? What alternative understanding of TF function is being put forward? The thrust of the argument here is that current structure-function mechanisms cannot but fail because they are inherently narrative and characterize molecular functions qualitatively: "Upon stimulation of the cell, TF-X translocates into the nucleus to bind promoter-Y and turn it on by recruiting cofactor-Z." A more principled conceptualization

of a TF's effect on gene regulation is that it has a finite occupancy at all DNA sites, scaled by their continuously varying affinities, and consequently there is a distribution of its associated catalytic activities over all DNA sites via similarly probabilistic PPIs with enzymatic cofactors. (Here I do not even address the competition for enzymatic cofactors, though that likely bears significantly on the system.) A TF (that is, the population of all the copies of a TF) applies something like a regulatory wave-function over the whole genome that interferes constructively and destructively with all the other wave-functions contributed by different TFs. When at a promoter the product of the probabilities contributed by various nuclear factors exceeds the threshold set by kinetic barriers, there is a response in the form of RNA synthesis. The well-documented "bursty" behavior of transcription reflects the probabilistic nature of this underlying regulatory phenomenon.[11]

Such a probabilistic conception of gene regulation should inform epigenetics, which is perhaps the prime example of how saturated molecular biology is with notions of structure yielding function—not just at the level of a biomolecule (as in protein folding) but fractally up the scale of complexity. In its fullest scope, the word *epigenetic* denotes the heritability of phenotype either through cellular propagation or organismal reproduction, and it has long been assumed that there must be a molecular structure underlying such heritability. The field found its structural platform in nucleosome-based chromatin, which seemed to be the literal realization of epigenetics as it is the complement of proteins upon (*epi*) DNA (the *genetic* material). So prevailing did that hypothesis become that *epigenetics* is now essentially synonymous with modifications to chromatin. Which is to say that *epigenetics is widely presumed to operate in cis*, at least by molecular biologists (systems biologists rely on regulatory networks to explain epigenetics).[12] However, processual reasoning can account for epigenetic phenomena as well: The spatial distributions, chemical states, and abundances of all the molecules in a cell are consequential of its history and causal of its future (along with input from without). The reason there is the appearance of memory or inheritance is because trans factors are transmitted through cell division and are capable of self-perpetuation through autoregulation. From this view, phenotypic durability is not like macroscopic durability that arises from a static structure, be it ink on paper or a post in the ground; rather, it is something like neuronal connections, which depend on ongoing stimulus (trans factor activity) to be maintained.

Is the cis environment therefore wholly irrelevant to epigenetics? Not at all, since it determines the kinetic barriers to regulatory processes. However, this does not imply that the cis environment is epigenetically instructive because trans factors regulate the cis environment.[13] True, some of the enzymatic activities concentrated at transcribed promoters by virtue of their TF occupancy are DNA- or nucleosome-modifying, which could provide some fleeting hysteresis. However, those activities are balanced by passive and active processes continually tending toward repression, such that in the absence of trans factors, there will be relaxation back to the default state.[14] The relevant regulatory or epigenetic question is what tips the (probabilistic) scale one way or another. Since trans factors operate upon the chromatin and not the other way around, the cis environment should be regarded as consequence, not cause, of regulatory signaling.

It could be objected that the cis environment is instructive because where trans factors are operative depends on things such as promoter accessibility. Such reasoning works only superficially, however, since once it is asked what determines the accessibility of a promoter, the instructive role of trans factors is again affirmed. To say that a promoter must be accessible for TF-X to bind it is tantamount to saying that TF-X depends on TF-Y to bind the promoter because whether it is

accessible or not is just a reflection of the presence or absence of TF-Y. Hence, the epigenetically relevant information is not the state of the promoter (in cis) but the presence of TFs (in trans), though the former can be used to report on the latter.

As an illustrative aside, "opening" DNA elements was once ascribed to a supposed subset of TFs dubbed "pioneer factors," but apropos of the previously discussed continuum of TF–DNA affinities, it was recently shown that "pioneering" can be done by any TF depending on its affinity for a given DNA element.[15] The role of the cis environment—especially the presence of generic DNA binders in the form of nucleosomes—is therefore to present a kinetic barrier by competing with trans factors, which *allows for* but does not of itself *constitute* regulation. For instance, when one TF displaces a nucleosome, another TF can bind the newly uncovered site, creating indirect cooperativity between the two TFs.[16] The nucleosome's position can in no way be construed as regulatory, however—only regulated.

It is still important for epigenetics that kinetic barriers be present in cis for, otherwise, all nuclear function would be an instantaneous reflection of the trans environment with no information of the previous state retained (except arguably through network evolution). The fact that promoter accessibility, for example, can be altered in a durable way means that the kinetic barrier to TF function can store information. One way in which that might work is the following: Imagine that the current complement of TFs in a nucleus would not be able to "pioneer" a given promoter but are able to maintain that promoter once it has been opened. That means if a factor had pioneered the promoter in a previous state, it could leave a lasting imprint—an epigenetic mark—in the form of promoter accessibility. That accessibility is not a static structure but a dynamic state that reflects the trans environment. The cis–trans interaction therefore allows for memory of the past and plasticity in the present.

All epigenetic phenomena can be abstractly understood as the self-perpetuation of trans factors as they interact with the cis structures that present kinetic barriers to their functions. At any given point in a cell's history, there will be a complement of TFs present. Those TFs in combination will have differential affinities for all promoters and for all transcription cofactors. The likelihood of a promoter that is not being transcribed to begin being transcribed is a function of all the TFs and coactivators present and the current cis configuration. Certain balances of TF concentrations (and extracellular contexts) are capable of self-perpetuation and define a cell state, while others are not and so continue to evolve until a stable configuration is reached. Hence, cellular states are at once heritable yet sensitive to perturbations of the trans environment.

Dramatic demonstration of that fact is provided by Yamanaka's success in inducing pluripotency in terminally differentiated fibroblasts by ectopic expression of but four TFs. Indeed, inducible pluripotency is perhaps the greatest demonstration of all this essay is trying to make explicit. Yamanaka's shocking results should have sufficed to dismantle the entire molecular biology paradigm. If there had been any doubt that the state of the cell is wholly determined by the complement of trans factors, it should have ceased as of August 2006.[17] Apart from trans factors being able to exert effects across the entire genome that are dominant with respect to the cis environment, how could a handful of TFs effect a reversal of development—the same TFs, in fact, as are necessary for stem-cell maintenance? Hence, the structure-function model of epigenetics fails.

But this profound result also runs completely counter to a network-evolution model of epigenetics, in which differentiation progresses by programmed changes from one stable attractor state to another. Yamanaka did not discover an evolved "program" which takes a fibroblast and turns it into a stem cell. Rather, the combination of the four Yamanaka factors is sufficient to reverse differentiation and erase nearly all memory of past experience. This is possible because cellular states are plastic—so plastic, in fact, that they are fluid. They are not identities but metastable cis–trans configurations which, with sufficient stimulus, can be completely altered. Such fluidity seems to be the basis of the phenotypic plasticity seen in T-cells[18,19] as well as in malignancy.[20–24]

Adopting the probabilistic view of protein–DNA and protein–protein interactions as well as the processual understanding of nuclear function that is its higher-order counterpart may at first seem impractical to the experimentalist. The scientific method of controlling variables to isolate and quantify individual influences on a system is inherently reductionistic; how can it cope with layers of continuous distributions underlying a literally fluid information-processing system? First, let it be pointed out that whatever the answer to that question is, it cannot be to pretend that reality is what it is not so as to hoodwink the scientific method. Second, never before has technology been so suited for dealing with systems of immense complexity. So-called artificial intelligence (AI) meets in the present day with rich, multidimensional, quantitative biological datasets—especially those derived from live-cell imaging, which preserve the dynamics which are not just part of the system but *are* the (processual) system. Neural nets in particular seem suited to the job of learning the small manifold contributions of a TF on all CREs, for instance. But regardless of how it is done, it is now incumbent upon the researcher not to turn a blind eye to the complexity of molecular biology but to capture as much of that complexity in as parsimonious a manner as possible. "Mechanism" may in many instances have to give way to prediction of input–output relationships, at least until theoretical and computational advances enable a degree of granularity currently obscured by intractable complexity.

There are also intuitive lessons to be gleaned from facing the messy molecular reality that is the nucleus head-on. The addiction of certain cancers to certain TFs is a salient example. Because a cell state depends on a certain composition of trans factors, a cancerous cell cannot exist, much less grow and divide, without certain TFs any more than a stem cell can remained undifferentiated without its "master regulators." In the case of breast and prostate cancers, it so happens that those TFs are ligandable through their hormone-binding domains, making such addictions easier both to identify and to exploit; but in principle, all cancer cells should have similar dependencies. Moreover, it might be possible to throw the transformed cell state off course by activating alternate pathways as much as by inhibiting those on which it actively depends. If the cell state can be destabilized, it is by no means guaranteed that it will readily find a new stable configuration, especially in the context of the whole organism and the surveillance of the immune system. Apart from cancer, it would be worthwhile to determine which of the trans factors present are necessary for any cell state—or which combination is sufficient—for self-perpetuation. From there, it would also be of both foundational and translational interest to map the paths of least resistance (fewest genetic or environmental changes) from one cell state to another.

To contemplate the nucleus without the simplifying but illusory intuitions derived from macroscopic experience is to appreciate how fluid its organization is. The discovery of the structure of DNA generated the field of molecular biology as we know it, and as a result, the framework of function arising from structure was strongly imprinted on generations of its practitioners. But while

it makes sense that the genetic information be encoded in highly structured molecule in order to persist through countless transmission events, it does not at all follow that the living, breathing, dynamic cell with all its fluctuating parts need abide by the same principle. To insist on structure–function relationships throughout the hierarchies of biology from DNA to cell, from cell to tissue, from tissue to organism, etc. would be to lose the living forest for so many embalmed trees. Switching metaphors, the DNA should not be conceptualized as an orchestral score dictating precise relationships between the different voices in the symphony. The music is far less determinate than that. The DNA does provide a necessary underlying structure—the song form—but the style of life is, in the final analysis, improvisational.

Chapter Two: A New Model of Enhancer Function

This chapter contains work carried out in collaboration. Sections of this chapter were previously published as:

Karr JP, Ferrie JJ, Tjian R, Darzacq X. The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer-promoter communication. Genes Dev. 2022 Jan 1;36(1-2):7-16.

Introduction

Gene regulation involves the interplay of genetically encoded circuitry with dynamic input from cellular signaling. *Cis*-regulatory elements (CREs) provide the circuitry, while transcription factors (TFs) transmit the signals. CREs' latent potential awaits realization by TFs, while TFs' trans-regulatory function depends on CREs to direct them to their target loci. The fundamental mechanism underlying this decoding of biochemical signals by TF-CRE interactions has been brought into question by recent data from live-cell microscopy experiments showing spatiotemporal dynamics at odds with current models. The unexpected results from such experiments have yet to be satisfactorily reconciled with the long-standing rules of CRE-promoter communication learned from decades of genetics, biochemistry, and genomics.

We know that CREs operate by a sequence of molecular interactions. By being enriched in TF recognition sequences, CREs recruit TFs via protein-DNA binding. CRE-bound TFs recruit other TFs as well as transcriptional cofactors—proteins or protein complexes that typically bear histone-modifying or nucleosome-remodeling enzymatic activities[25]—via protein-protein binding. CREs thereby assemble a combination of proteins and enzymes at a particular position on the chromosome, while TFs translate DNA sequence into local enzymatic activity (acetylation, phosphorylation, methylation, etc.) via their DNA-binding and protein-interaction domains, ultimately regulating RNA polymerase II activity at a target promoter.

How such interactions at a proximal CRE could regulate transcription is conceptually much more straightforward than at a distal CRE, which can be many kilobases, or even a megabase, upstream or downstream from the target gene. There are at least three conceivable models by which a distal CRE could operate (Figure 2.1). **Model 1—"Stable-contact model"**: By the formation of a long-lasting protein-DNA complex stabilizing a chromatin loop, the distal CRE and promoter effectively become a single compound CRE with properties that neither element possesses on its own. **Model 2—"Kiss-and-run model"**: By transiently contacting the promoter, the CRE could deposit some material onto the promoter—be it TFs, other components of the transcriptional machinery, or post-translational modifications (PTMs) of promoter-bound proteins—which persist beyond a transient CRE-promoter contact. **Model 3—"Communication by diffusion model"**: The CRE could communicate with the promoter in a distance-dependent manner through the diffusion of TFs activated by enzymes recruited to the CRE.

Each model has distinct requirements and temporal predictions. Both Models 1 and 2 fundamentally require direct contact between the CRE and promoter via DNA looping (note: by direct contact we mean an unbroken chain of molecular binding interactions, which we expect not to exceed 10s of nanometers). However, whereas Model 1 proposes that the promoter is active only when in contact with enhancer and therefore necessitates persistent DNA-protein complexes tethering the CRE to the promoter, Model 2 allows some memory of interaction and thus requires only transient CRE-promoter contacts at some frequency. Model 3, on the other hand, does not

necessitate contact between CRE and promoter, but does have a quantitative dependence on proximity and therefore predicts that sustained contact would strengthen the effect of the CRE on the promoter.

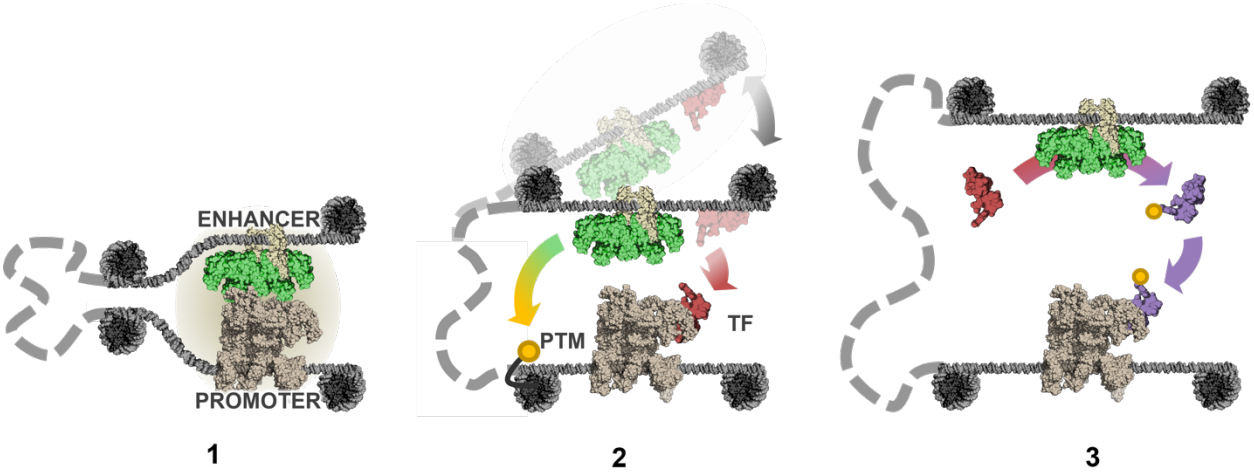

Figure 2.1—Mechanisms of Enhancer-Promoter Communication: There are three different ways in which a distal enhancer could regulate a promoter. (1) Stable contact model: A "compound cis-regulatory" element is formed by a stable complex of TFs (tan), coactivators (green), and the transcriptional machinery (grey). (2) Kiss-and-run model: Upon transient contact, enhancer-bound coactivators deposit PTMs at the promoter and TFs (red) are transferred. (3) Communication by diffusion model: TFs are activated (purple) at the enhancer and diffuse to the promoter.

Model 1 hearkens back to bacterial gene regulation and offers an intuitive solution to the problem of distance between enhancer (a representative class of CRE) and promoter, making it the prevailing textbook picture of *cis* regulation. It further gained popularity as the previous decade saw marked advances in chromosome-conformation capture (3C) technologies, which have detected signal enrichments between enhancers and promoters.[26,27] Although such 3C signal has been widely interpreted to indicate contact, it should be remembered that 3C does not actually report on contact between genomic regions in a single live cell, but rather the probability of cross-ligation in a large population of fixed cells. That is to say, it reflects not temporal frequency but population frequency, captured under chemical crosslinking conditions, which are known to perturb both chromatin structure and TF-chromatin interactions.[28,29] Hence, 3C data cannot be used to test different temporal predictions about the longevity of CRE-promoter contacts in vivo. Further, 3C methods cannot discern whether CRE-promoter contact is at all necessary for transcription in single cells, so they cannot verify or falsify any of the three models.

To test temporal predictions, microscopy experiments are required to fill in the gaps in our knowledge concerning CRE-TF interaction dynamics and CRE-promoter distances in single cells. Such experiments have consistently surprised us by yielding results dissonant with expectations from Models 1 and 2, and even incompatible with their basic requirements. Single-particle tracking experiments measuring the diffusion of nuclear proteins have documented fleeting lifetimes of TF-chromatin interactions, from hundreds of milliseconds to several seconds.[30] Such rapid dynamics are obviously difficult to reconcile with Model 1. Moreover, recent experiments measuring both

distal CRE-promoter distances and promoter activity in single cells[31] have not supported either Model 1 or 2. For enhancers removed by scores to hundreds of kilobases, promoter activity was shown to have dependence on proximity (~350 nm) in one case,[32] an anti-correlation with proximity in another case,[33] and no correlation with closer proximity in another case.[34] A FISH method able to probe many interactions in a single cell likewise saw weak correlation between contact and activity in some enhancer-promoter pairs, and no correlation in others.[35] The basic requirement for contact in Models 1 and 2 is not satisfied in these instances. Moreover, several experiments perturbing proteins involved in global 3D genome organization have shown at most mild effects on transcription output despite profound losses of 3C signal.[36–39] It is theoretically possible to reconcile the kiss-and-run model with such weak or absent temporal correlations of promoter activity and enhancer-promoter proximity only if it is assumed that there is an enduring "memory" of interaction such that each contact contributes to an accumulating signal at the promoter, whether in the form of proteins or of protein PTMs.[40] We find that to be a dubious assumption given the transience of protein-DNA interactions and of PTM lifetimes and note that it still strictly requires contact; indeed, the frequency of contact must be inversely proportional to the length of memory. Last, because the modeling from Xiao et al. (ref. 35) was done in arbitrary time, we cannot know whether the infrequency of contacts seen in recent microscopy experiments is reconcilable with the proposed theory.

We are led then to one of two conclusions: either the recent studies from multiple groups were technically unable to observe the phenomenon of enhancer-promoter looping in single cells, or it is much rarer and of lesser regulatory importance than has been supposed. Therefore, although contact-dependent models have not been entirely disproven, to consider an alternative model may nonetheless be warranted by the new evidence at hand and could prove useful in instigating discussion of a broader range of mechanisms. The communication by diffusion model has largely been disregarded because it poses a fundamental problem believed to be irreconcilable with the physics of diffusion—namely, it depends on TF molecules visiting the CRE and subsequently binding promoters in *cis* with a higher probability than other DNA elements in the nucleus. However, we have conceived of what seems to be a plausible mechanism for CRE-promoter communication via diffusion. We call it the TF-activity gradient (TAG) model, since it consists of CRE-associated **enzymes** modifying **TFs** to create **local 3D gradients of chemical signals**. We find the TAG model attractive in that it derives naturally from longstanding but previously unconsolidated observations, and it grounds CRE-promoter communication in exquisitely regulable enzymology, without relying on the more topologically constrained and convoluted process of intrachromosomal contact.

A New Model of CRE-promoter Communication

The road to the TAG model began with recognizing that the substrate ranges of so-called histone-modifying enzymes (or "epigenetic writers") are actually not restricted to histone substrates, but invariably include TFs.[25,41,42] Hence, positions along the chromosome enriched for histone PTMs represent likely sites for enhanced TF modification. This suggests the intriguing possibility that CREs, which bear a significant number and diversity of histone modifications, could act as inducible platforms for catalytic modification of TFs. CREs could thereby serve a function analogous to that of scaffold and targeting proteins for signaling kinases—namely, to bring together promiscuous enzymes with specific substrates in a regulable and localized manner, or to coordinate signal relays by clustering different enzymes in a pathway together.[43]

Take for example the transcriptional coactivator and lysine acetyltransferase CBP/p300 (henceforth just p300), which has long been appreciated as a central player in gene regulation. Levels of histone H3K27 acetylation, its signature chromatin mark, at CREs have been used to predict nearby promoter activity[44,45] and recruiting the catalytic core of p300 to enhancers via a dCas9 fusion is sufficient to activate target promoters.[46] However, the precise mechanism by which p300 regulates transcription has remained unclear. Some puzzling reports have indicated that although p300 catalytic activity is necessary for enhancer function,[47] H3K27 acetylation is not.[48,49] This discrepancy could be resolved if non-histone substrates were its functional targets and the histone mark only a collateral effect of its local activity. Since H3K27ac has never been causally linked to transcription and p300 does indeed exhibit a promiscuous substrate range, including scores of TFs whose regulatory activities are often modulated by acetylation[50,51],[50,51] we entertain and expand on this possibility here. (It should be noted that p300 was dubbed a histone acetyltransferase only because of the historical coincidence that histones were its first discovered substrates, not because they were demonstrated to be its specific or functional substrates.)

If TFs are acetylated at p300-bound enhancers (Figure 2.2), the result will be spatially heterogeneous distributions of chemical signals in the form of acetylated TFs (ac-TFs). To see this, let us consider what happens immediately after a TF is acetylated by p300. As time passes, the further the ac-TF diffuses from p300, and the more likely it is to encounter a deacetylase—an abundant and ubiquitous class of enzymes also named after their histone substrates ("HDACs") even though they have many others.[52] As a result, p300 at a CRE becomes the point-source of a concentration gradient of ac-TF. If the point source is free to diffuse throughout the nucleus, local concentration gradients will not form. However, if active p300 is bound to chromatin while the TFs remain diffusible, then a gradient will arise centered on the enzyme-bound chromatin region—i.e. an enhancer. Consequently, a promoter proximal to an enhancer is far more likely to encounter an ac-TF than a promoter distant in 3D space.

Note that at equilibrium—in the absence of PTM deposition—there cannot be stable gradients of TFs arising from CREs. The existence of a nearby binding site for a TF in no way enhances its equilibrium occupancy at the promoter (in fact, the more competing sites there are, the lesser the occupancy will be at a given promoter). Even if we consider that the two elements contact one another at some rate (as in Model 2), the equilibrium remains unchanged by the contact. Although it is true that upon contact a CRE-bound TF would have a higher likelihood of unbinding the CRE and binding the promoter, reciprocally a promoter-bound TF would be more likely to unbind the promoter and bind the CRE. Without invoking non-equilibrium processes, therefore, the only way in which a distal binding site could increase the TF's occupancy at the promoter is if simultaneous binding at both sites through DNA looping cooperatively strengthened the TF-promoter interaction (as in Model 1).

It is the presence of an enzymatically generated gradient of activity that overcomes the major difficulty with Model 3 alluded to before: TAG requires proximity (acts in *cis*) but does not depend on contact between enhancer and promoter (can operate over hundreds of nanometers in 3D space, accounting for very distal CREs). Enzymatic deposition of PTMs on TFs at a chromatin site would give rise to such a gradient for two reasons: (1) the volume through which the signal spreads increases cubically with radial distance from the enhancer, rapidly diluting the concentration of ac-TFs, and (2) an approximately constant rate of deacetylation due to abundant HDACs leads to an exponential decay profile at steady state (similar to morphogen gradients). The *cis* relationship

between enhancer and promoter is therefore maintained not by the two *cis* elements physically associating, but by the biophysical and biochemical limitations on the extent of the signal diffusing from the enhancer.

Unlike the kiss-and-run model, TAG does not rely on a coordinated, vectorial transfer of material from enhancer to promoter; it is mediated by the random motions of ac-TFs emanating from the enhancer, with a small subset finding the promoter. TAG therefore has an inherent inefficiency: since each ac-TF has a low probability of finding its target promoter, enhancer-bound p300 must modify many TFs over time. This inefficiency is "paid for" not in the currency of acetyl-CoA, which gets regenerated in situ by nuclear synthases, but in the ATP those synthases consume[53].[53] The probability of the ac-TF finding its target can also be increased in two ways: by altering the mode of its diffusion or by effectively increasing its target's size.[54] If the ac-TF undergoes sub-diffusion which causes it to more densely sample its immediate environment—for instance, by sliding or hopping on chromatin—it becomes much more likely to find a local target[55].[55] If at the target promoter there is even a transient hub of locally concentrated proteins (as have been observed in various systems[56–61]J. Chen et al. 2014; W.-K. Cho et al. 2018; Chong et al. 2018; Z. Liu et al. 2014; Mir et al. 2017; Tsai et al. 2017)) with affinity for the ac-TF, this could multiply the probability of each TF finding the promoter and thereby sensitize the promoter to the enhancer. Isotropic diffusion from the enhancer would result in a broad (if confined) range of enhancer influence and thereby allows for one enhancer to regulate multiple promoters—which is a documented property of enhancers.[62,63]

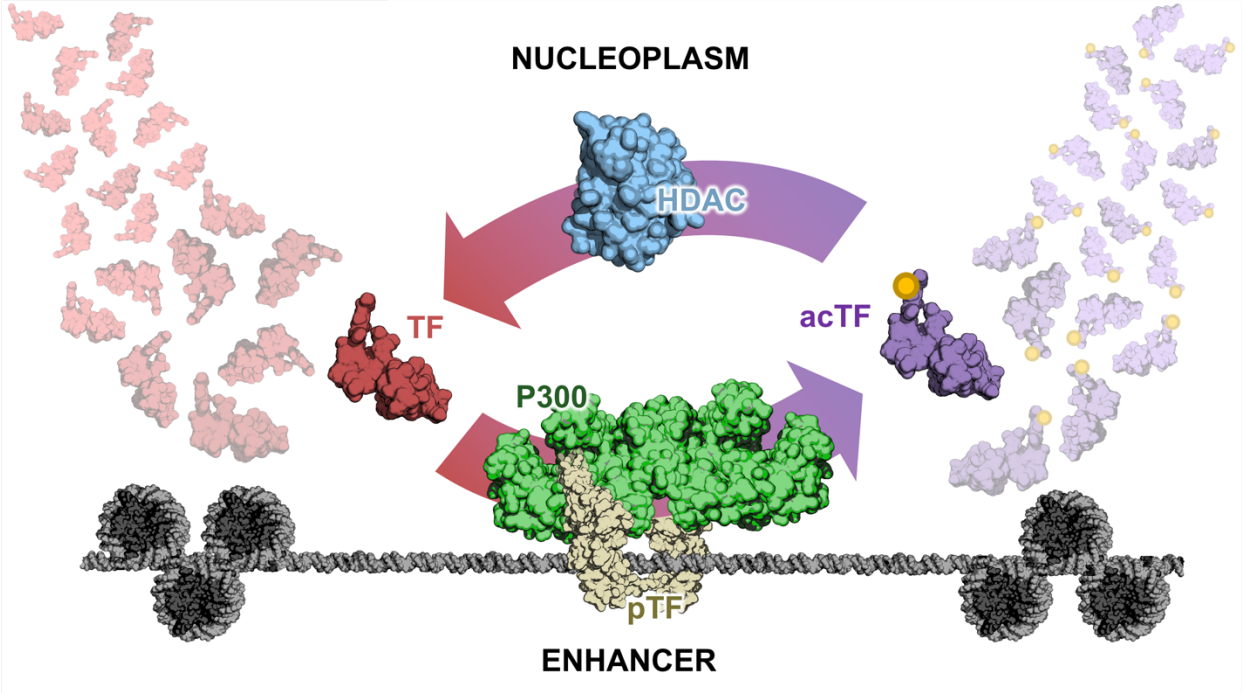

Figure 2.2—Picturing the enhancer as a point-source of acetylated TFs. An unmodified TF (red) contacts an enhancer bearing activated p300 (green) recruited by a dimeric transcription factor (tan). The acetylated transcription factor (purple) departs from the enhancer and is recycled by being rapidly deacetylated in the nucleoplasm by abundant histone deacetylases (blue).

Based on published findings, we propose the following TAG-based mechanism for p300-mediated regulation of transcription occurring at a generic enhancer-promoter pair: (1) p300 is first recruited to the enhancer by a sequence-specific DNA-binding TF; (2) an allosteric regulator binds and activates p300; (3) active p300 acetylates nearby substrates, including both histones and TFs bound to the enhancer; (4) subsequently acetylated TFs (with typical residence times of a few seconds) disengage from the enhancer and diffuse outward; (5) when ac-TFs reach a target promoter, they increase transcriptional output; (6) diffusing ac-TFs are deacetylated at a high rate, limiting the spatial range of their action. We will now walk through this hypothetical realization of the TAG model and present some supporting evidence for each step while pointing out some of the unknowns.

(1) p300 is recruited to the enhancer

p300 is such a pervasive enhancer-bound coactivator that its ChIP-seq binding profile is routinely used to identify enhancers.[64,65] Not much is known about the recruitment mechanism of p300, except that it depends on sequence-specific TFs and is subject to competition, as the number of p300 molecules per cell (approximately 7000) is on par with the number of active enhancers.[66,67] Thus, with respect to p300, the nucleus is indeed a heterogeneous landscape, with enhancers being rare landing pads to enable and direct its enzymatic activity.

(2) An allosteric regulator of p300 binds the enhancer and activates p300

The presence of H3K27ac, and not p300, differentiates active from poised enhancers,[68,69] indicating that at least in some instances recruitment and activation of the enzyme are separable events. Biochemical and structural studies have uncovered allosteric regulation of the catalytic activity of p300. Similarly to how some kinases dimerize and undergo trans-autophosphorylation, p300 is activated by trans-autoacetylation.[70] Dimeric (and often phosphorylated) TFs mediate this trans-autoacetylation by bringing together two p300 molecules.[71] This mechanism integrates cellular signaling in that many cytosolic signaling pathways lead to dimerization and nuclear translocation of TFs to affect gene expression.[72] For many inducible TFs, signaling-mediated oligomerization has been shown to recruit active p300 at target enhancers and promoters.[73–80] Where these TF-p300 complexes bind is cell-type specific and depends on the prior establishment of accessible CREs by pioneer factors.[81]

(3) Active p300 acetylates TFs bound at the enhancer

To our knowledge, it has never been investigated whether p300 acetylates non-histone substrates on chromatin or throughout the nucleoplasm. However, it was found that p300 substrate specificity is higher in vivo than in vitro,[51] indicating that its cellular context is important for restraining the action of this highly promiscuous enzyme. Additionally, the fact that known TF substrates of p300 have ChIP-seq signals at chromatin loci enriched for H3K27ac suggests that at least some TF acetylation may occur on select chromatin sites. Furthermore, for many TFs, the dimerization that is required for p300 activation is typically required for their DNA binding as well, with nuclear receptors being an iconic example.[82] Furthermore, in the case of the TF p53, consensus-sequence DNA acts as an allosteric ligand to promote acetylation by p300 through exposure of the acetylation motif,[83,84] suggesting that p300 substrate specificity can also be mediated by the proximal DNA. Multiple lines of evidence therefore point to chromatin as a likely site of p300's TF-modifying activity.

(4) Acetylated TFs disengage from the enhancer and diffuse outward

Within seconds of landing at a CRE and possibly being acetylated, a TF will unbind and continue diffusing.[30] Some acetyl-lysines are bound by bromodomains,[85] which can, in turn, change TF-chromatin interactions and thus alter TF occupancy at target sites.[86] Other acetyl-lysines induce conformational changes to expose new surfaces for interaction or (de)stabilize existing ones.[87] Acetylation can also weaken protein-DNA interactions by neutralizing the positive charge on a phosphate-interacting lysine.[88] Consequently, it is plausible that an ac-TF will exhibit distinct modes of diffusion from those of the unmodified TF.

(5) When acetylated TFs reach a target promoter, they increase transcriptional output

Increased transcriptional activity upon acetylation has been demonstrated for various TFs,[89–99] although the mechanism is not clear in most cases. One notable mechanism is through interaction with bromodomain-containing protein BRD4 to recruit P-TEFb to promoters and phosphorylate Pol II.[100] However, it is possible that the mechanisms at play are as diverse as the TFs being acetylated. Acetylation, like many PTMs, need not have one outcome but rather provides a regulatory switch that can modulate protein activity positively or negatively. Such flexibility may be key to achieving combinatorial specificity and complexity, allowing the same enzyme to exert positive or negative regulatory effects on transcription depending on what substrates are present.

(6) Acetylated TFs are deacetylated at a high rate, limiting the spatial range of their action

Once modified at a CRE, an ac-TF has a lifetime dictated by the abundance and activity of deacetylases, of which there are 18 varieties in humans.[101] Recently, a proteomics study of erythropoiesis documented what the authors described as a "vast quantitative imbalance" between the number of HDAC molecules—in the hundreds of thousands—compared to that of p300, at less than ten thousand.[67] Although live-cell imaging of HDACs is lacking, by immunofluorescence these HDACs are predominantly nuclear and homogeneously dispersed.[102] Evidently, the mammalian nucleus has evolved to keep global acetylation levels tightly controlled, which we speculate is to enable local signaling. Hence, we predict the extent of ac-TF signal to be exquisitely spatially restricted around the CRE point source.

Conclusions and Outlook

The TAG model offers a novel solution to the two key problems of CRE-promoter communication: what signal is transmitted from distal CREs to promoters, and by what mechanism does the transmission occur? We envision that the signal could be cofactor-deposited PTMs on TFs that alter their trans-activating/repressing potential, and the mechanism of communication is an enzymatically time-limited diffusion from the CRE point source. That TFs will be modified at CREs is a logical consequence of two established observations: CREs are TF-binding hotspots replete with histone signatures of cofactor activity, and those TFs are known substrates of cofactor enzymes. Chromatin-associated enzymatic activities, in turn, result in spatial heterogeneity of chemical signals given diffusible substrates and the profusion of de-modifying enzymes that limit the lifetime of the PTM. The requirement for distal CRE-promoter proximity is therefore met without necessitating direct or sustained contact between the two elements, and the parameter that tunes promoter activity is the PTM-TF flux from nearby CREs, not the frequency or stability of direct contact between two chromosomal regions.

Although it has long been known that CREs are hotspots of cofactor activity, the chemical ramifications of PTM-depositing enzymes being recruited to defined positions along the chromosome have largely gone unexplored. Perhaps this is partly due to the fact that TF-CRE binding interactions are widely treated as mere equilibrium-driven associations, instead of a platform for non-equilibrium modification of the TFs by CRE-associated enzymes—even though it has been appreciated since almost the beginning of the enhancer field that covalent modification of TFs was key to their regulation.[103] Given that PTM-depositing enzymes are enriched at CREs, where we now know TFs rapidly bind and unbind within seconds, and that so-called histone-modifying enzymes are vastly outnumbered by their corresponding de-modifying enzymes,[67] we propose that CREs function as local "reactors" to generate tunable concentration gradients of modified and activated TFs that diffuse to nearby binding sites to regulate target promoters.

Since PTMs have profound effects on TF interactions and function,[104] the ability to locally and transiently concentrate them provides a mechanism for precise control of gene regulation. If, for instance, a TF is only 10% active in its unmodified state, an increase in its nuclear concentration will have relatively mild effects on target gene transcription except in regions where it acquires its activating PTM. Consequently, TF-responsive promoters with nearby CREs able to deposit the PTM will be much more activated than target promoters lacking such CREs. Conversely, a TF may activate transcription when unmodified but repress transcription when modified. The ensemble of active CREs and their associated enzymes would therefore determine the global transcriptional changes in response to the presence of a TF. Consistent with this notion, it has been shown that inducible TFs bind pre-existing CREs that are made accessible by lineage-determining TFs,[105] that proximal promoter activity is correlated with recruitment of enzymatic coactivators to such pre-existing CREs,[81] and that DNA accessibility of CREs is just as determinative of TF occupancy as the presence of TF recognition sequences.[106]

Different TFs therefore play different roles in the TAG model. Certain TFs are necessary to define which CREs are accessible, other TFs bind the accessible CREs to recruit cofactors, and perhaps still other TFs are the substrates of those cofactors that will carry the signal to nearby promoters. Importantly, specificity is attained by sequence-specific protein-DNA and selective protein-protein interactions at each of these steps. So-called master regulators or pioneer factors find recognition sequences to open up a subset of CREs;[107] which TFs bind available CREs is determined by their specific DNA affinity;[108] and what proximal genes respond to an active CRE depends on the PTM-TF binding at the promoter—whether indirectly through protein-protein interactions with promoter-bound TFs or directly through the PTM-TF binding the DNA. Other mechanisms can also be at play due to the number of tunable parameters in the system. For example, disruption of TAD boundaries can, in some rare instances, lead to ectopic expression of a promoter,[109] indicating that at certain loci in specific cell types, part of the CRE-promoter specificity can be influenced by the 3D organization of local chromatin. Hence, although the TAG model is not explicitly concerned with the question of specificity—that is, why a CRE affects certain promoters in its vicinity and not others—it nevertheless hints at a possible path to specificity mediated, satisfyingly, by sequence-specific DNA-binding proteins and their protein-protein interactions.

Although it is conceptually helpful to imagine one PTM-TF communicating from a CRE to a promoter, it is likely that multiple protein species contribute to the PTM gradient, some of which may not be sequence-specific TFs. Because some active enhancers are transcribed, we can surmise that much of the transcriptional machinery is being recruited to them; components of that

machinery that are modified should also give rise to a local gradient due to diffusion from the CRE and deactivation by repressing enzymes. For instance, general transcription factors, the elongation factor P-TEFb, and RNA polymerase II (Pol II) are all acetylated by p300.[110–112] Of these, P-TEFb is an especially intriguing candidate in that it is broadly required for Pol II transcription, depends on p300 for activation, exhibits sub-diffusion, and can be rapidly inactivated by the 7SK complex.[54,112,113] The TF c-Myc may also be an important CRE-promoter relay molecule, in that it generally amplifies transcription from active promoters,[114,115] and is known to be acetylated and activated by p300.[116,117]

Since TAG depends on PTM of diffusible molecules, what role, if any, do histone modifications play in this model? It should be acknowledged that the respective contributions of histone PTMs and TF PTMs to gene regulation are challenging to disentangle—the fact that the same enzymes are likely modifying both substrates at the same locations makes correlations abound. However, it is difficult to imagine how PTMs at a CRE could participate in regulation of a distal promoter. It would seem that in order for histone modifications to be directly involved in CRE-promoter communication, a CRE-bound enzyme would need to modify promoter-bound histones, necessitating at least transient contact between these two elements. Even if it is granted that CRE-promoter contact is a requisite for regulation, there is no obvious mechanism for achieving specificity. If an active CRE-bound enzyme modifies any histones it contacts, how are promoters selectively modified over intergenic regions, or some promoters activated while others in the vicinity are not? Moreover, as noted above, the histone PTMs would need to have a lifetime much greater than that of the CRE-promoter contact, which is difficult to imagine given the pervasiveness of HDACs. This difficulty could perhaps be surmounted if a CRE-promoter contact initiates a feed-forward loop in which a histone mark deposited by the CRE-bound enzymes recruits more histone-modifying enzymes to the promoter. However, such a system seems alarmingly unregulable as well as unspecific. We therefore posit that histone modifications may not play an instructive or causal role in transcription initiation, but may rather have a permissive role in maintaining CREs in a particular state—for instance, by stabilizing TF-recruited and -activated enzymatic cofactors on the chromatin, or creating local chromatin landscapes that influence TF-chromatin interactions and therefore TF diffusion.

Another aspect of CRE-promoter communication that has garnered much attention for decades but has received little comment in this review is the role of boundary elements or insulators. These elements are described by three major features: in an orientation-sensitive manner they can prevent an enhancer from communicating with a promoter, they delineate boundaries between active and repressive chromatin marks, and they are associated with TAD boundaries in 3C assays. TADs have also been shown to correspond to domains over which an enhancer can activate a promoter.[63,118] Whereas Drosophila have several known insulator-binding proteins, the only one that has been characterized in mammalian cells is CTCF. Since depletion of CTCF has very mild effects on transcription, both in number of genes affected and in the magnitude of that effect model accounts for boundary elements in that TAD organization will inform 3D distances, which will then affect the ability of a CRE to influence a promoter—that is, if a topological boundary moves a promoter outside an enhancer's gradient of activity, it should effectively insulate the two elements in a seemingly stepwise fashion. Within a TAD, however, TAG would predict that there would be a graded effect as a function of enhancer-promoter distance, as has been documented at least once.[119] The less compact the chromatin is, the more dramatic should be this effect, which could explain why the loop-extruding cohesin complex is necessary for distal but not proximal

enhancers.[120] The insulator-defined boundaries of histone modifications also suggests that between TADs, protein-chromatin interactions may differ and could change the local diffusion dynamics of TFs (e.g. if one TAD has more accessible DNA than a neighboring TAD and thereby better retains the TF[121]).

Various groups, including our own, have observed that at the sub-TAD level, a remarkable pattern of enhancer-promoter cross-ligation is visible by Hi-C and Micro-C.[27,122] Additionally, the Engreitz lab was able to use Hi-C signal enrichments in a powerful strategy to predict functional enhancer-promoter pairs.[45] Their model includes only a few parameters: read counts of H3K27ac ChIP-seq and DNase hypersensitivity—which they collectively dub "activity"—and "contact" (Hi-C enrichment). Impressively, the product of activity and contact normalized against the surrounding 5 Mb region correlated with quantitative effects on gene expression with 70% precision at 70% recall. The authors further make an intriguing observation that Hi-C contacts can be replaced by linear distance with little damage to the power of the model. The implication of this is that higher levels of enzymatic activity can compensate for greater distances between enhancers and promoters—a result that is easily rationalized by TAG: at greater distances from the enhancer, the greater the fold-decrease in the ac-TF gradient, so the greater the initial signal (activity) must be to compensate. We therefore propose that a more conservative interpretation of 3C data may also be the more biologically relevant: 3C signal may actually reflect 3D proximity, not contact. We also suggest that the ability of higher activity to compensate for greater distance is more easily explained by TAG than by contact-dependent models.

Direct demonstration of the TAG model would require tracking locally deposited and exceedingly transient chemical modifications of diffusing proteins in live cells. Although local gradients of a small molecule in vivo have been measured,[123] detecting a PTM gradient would require new technological developments. Nevertheless, TAG makes some predictions that differentiate it from other models. First, CRE-promoter communication should be dependent on distance but not on contact. Unfortunately, with the spatiotemporal resolutions of techniques currently available, these two parameters are difficult to tease apart. Second, TAG predicts that the regulatory effect of a cofactor—whether activating or repressive—will depend on the TF modified. This prediction differentiates TAG from histone-centric models in that a histone modification would be predicted to have the same effect at different CREs (unless, as the histone code hypothesis proposed, histone PTMs occur in complex combinations which have emergent regulatory properties; but the high redundancy of histone marks suggests this mechanism is not likely widely employed[124]). Furthermore, TAG predicts that gain-of-function mutations of cofactor-modified residues of TFs should bypass the dependence on CREs at target promoters where the PTM-TF is sufficient for activation or repression.

While it requires substantial experimental validation, the model presented here provides one plausible alternative to DNA-looping models for how TF inputs are dynamically processed by genetic circuitry into transcriptional outputs: DNA sequence determines local enzymatic activity, which in turn dictates the regulatory function of diffusible TFs at proximal promoters. Such a mechanism harnesses the properties of the nucleus that make it qualitatively different from a test tube—spatial confinement and heterogeneous distributions of molecules—allowing local effects and not just global parameters (e.g. TF concentration) to have profound regulatory impact. While posing challenges to current experimental techniques, such a system would afford exquisite control of gene expression via precise DNA-targeted enzymatic control of promoter microenvironments.

We conjecture that CREs evolved to do just that—to generate neighborhoods of chemical signals in which TFs can play TAG.

Chapter Three: p300 Is an Obligate Integrator of Combinatorial Transcription Factor Inputs

This chapter contains work carried out in collaboration. Sections of this chapter were previously published as:

Ferrie JJ*, Karr JP*, Graham TGW, Dailey GM, Zhang G, Tjian R, Darzacq X. p300 Is an Obligate Integrator of Combinatorial Transcription Factor Inputs. Molecular Cell, In Press. *these authors contributed equally

Introduction

Eukaryotic transcription factors (TFs) depend on coactivators to activate transcription. Coactivators have no intrinsic ability to bind specific DNA sequences, and their recruitment to specific target loci is instead proposed to rely on interactions with TF activation domains, modified histone tails, and other coactivators. Accordingly, most coactivators are multi-domain or multi-subunit complexes containing both enzymatic domains and diverse protein-protein interaction modules, including epigenetic reader domains (e.g., bromodomain, PHD domain), TF-interaction domains (e.g., KIX, PAS), and intrinsically disordered regions (IDRs).[125–127] The relative importance of these various domains for target binding remains unclear. Moreover, it is puzzling how coactivators get distributed among target loci rapidly and reproducibly despite being stoichiometrically limiting relative to TFs and even to active cis-regulatory elements.[66,67,128]

To clarify these outstanding uncertainties in the field, we focused on p300—a central node in gene regulation that combines many domains in one polypeptide, making it more genetically tractable than multi-subunit coactivator complexes (e.g., Mediator). p300 is composed of three large regions: the N-terminal region (NTR), a well-structured enzymatic central region (Core), and the C-terminal region (CTR) (Figure 3.1A). In the Core are several annotated chromatin-interaction domains (ChIDs), while interspersed through the highly disordered NTR and CTR are various TF-interaction domains (TFIDs)—small helical and zinc finger domains that have been shown to engage in "fuzzy" binding with TF activation domains.[8]

There are conflicting reports in the literature of how p300 interacts with chromatin. In vitro binding assays have long shown that the various TFIDs distributed throughout p300 bind to myriad peptides contained within disordered activation domains of TFs.[129] Many years of such biochemical data support the model that TF binding at a locus leads to local enrichment of p300 through TF–TFID interactions. However, other biochemical data suggest that p300's bromodomain binds to acetylated histone tails, providing a distinct mechanism for p300 chromatin targeting in line with the histone code hypothesis. Indeed, more recent live-cell experiments have suggested that p300 binds to acetylated lysine residues, including those produced by p300's own acetyltransferase activity, and cryo-EM structures of the interaction between p300 and acetylated histone tails have recently been published.[130–133] Last, the widespread enthusiasm for phase separation as an organizational principle of the cell has led to the hypothesis that the extensive IDRs present throughout p300 could provide a driving force for p300–TF interactions and downstream regulation of transcription.[132,134] These three distinct though not necessarily mutually exclusive mechanisms of p300 interaction with chromatin have not been directly tested head-to-head in a living cell, leaving ambiguity in the literature concerning a major gene-regulatory step.

We therefore set out to address two questions: First, do ChIDs, TFIDs, or IDRs—or some combination thereof—mediate p300 chromatin engagement? Our hypothesis was that the TFIDs

provide sequence specificity to direct p300 to target sites while the ChIDs bind to histones to stabilize p300 on chromatin. By recognizing particular histone marks, the ChIDs could also contribute to directing p300 to target sites bearing active marks.[135] Others have also argued that the IDRs of p300 cause it to partition into TF condensates.[136] The second question is whether there is interplay between p300 chromatin binding and catalytic activity. Because the p300 Core has both an acetyltransferase domain and a bromodomain—which is thought to bind acetylated lysines, especially on histones[137]—we considered that enzymatic activity might enhance chromatin binding and stabilize active p300 at its target sites.

Results

To address these questions, we turned to high-speed (4 ms/frame) single-molecule tracking (SMT) combined with a Bayesian analysis method that infers underlying diffusive states from a population of observed trajectories.[138] The output of this analysis is a "diffusion spectrum," or probability density for every value of diffusion coefficient (Figure 3.1B). The fractional likelihood of a molecule moving with a diffusion coefficient of <0.1 μm2s-1 we call the bound fraction (f_{bound}) because it represents the proportion of molecules diffusing at a rate indistinguishable from that of chromatin motion (see Halo-H2B in Figure 3.1B). By measuring their effects on f_{bound}, we are thus able to assess the impact of various p300 mutations and perturbations on chromatin interactions in the context of the live cell.[138,139] Although the rest of the diffusion spectrum faster than 0.1 μm2s-1 holds other valuable information, such as how many distinct diffusive species (number of peaks) exist for a given tagged protein, we have focused on changes to f_{bound} in this report.

We first benchmarked SMT behavior of the gene products of two HaloTag knock-ins at the endogenous EP300 locus against stably integrated Halo-H2B and Halo-NLS transgenes in U2OS cells (Figure 3.1B-C). As expected, Halo-H2B was predominantly bound (f_{bound} = 0.85 ± 0.01) while Halo-NLS was predominantly fast-diffusing. (We note that Halo-NLS has an f_{bound} of 0.11 ± 0.03, which we consider the baseline of the assay.) In this cell line, p300 had an f_{bound} of approximately 0.54, which was reproducible between two different clonal knock-in lines (f_{bound} = 0.55 ± 0.02, 0.52 ± 0.03). We also used FRAP to measure the residence time of each p300 construct on chromatin and found that in contrast to H2B, which has a residence time far beyond the timescale of this experiment, p300–chromatin binding events persist for approximately 26 seconds on average—on par with or slightly longer than residence times typical of TFs.[140]

Having established a characteristic profile for WT endogenous p300 diffusion, we built a transgene system to facilitate mutation of p300. In order to avoid complexities arising from interactions with endogenous p300, we generated a clonal knock-out cell line expressing no detectable p300 (Figure 3.S1A) into which various p300 transgenes were introduced by random integration and antibiotic selection. The full-length p300 transgene product behaved similarly to the tagged endogenous protein (f_{bound} = 0.51 ± 0.02—Figure 3.1B-C), validating this assay system. The stable transgene was considerably less expressed than the endogenous protein, but we verified that in our system f_{bound} is not sensitive to concentration of protein (Figure 3.S2).

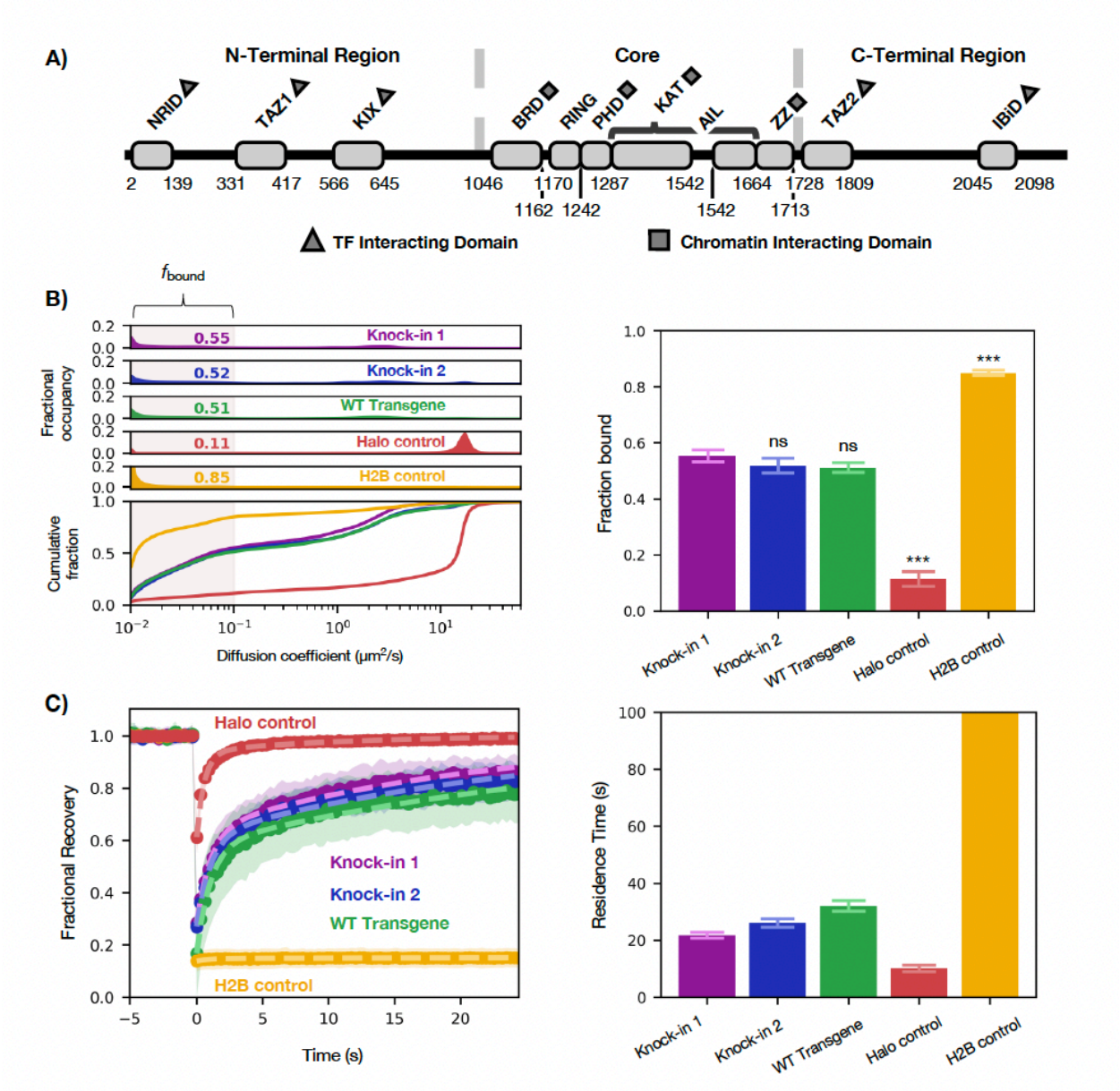

Figure 3.1—Benchmarking Halo-p300 diffusive behavior.
(A) Schematic of domain organization of p300 showing the TFIDs in the NTR and CTR and ChIDs in the Core. The black line indicates IDRs. Relative lengths are not to scale. (B) Diffusive spectra (left)—probability density function (top) and cumulative distribution function (bottom)—with plot of f_{bound} (right). Bars represent bootstrapping mean ± SD while p-values are reported as ns, (not significant), * (< 0.05), ** (<0.01), or *** (<0.001). (C) FRAP curves of the same (left) with residence times plotted for each construct (right). Bars represent best fit ± 95% CI.

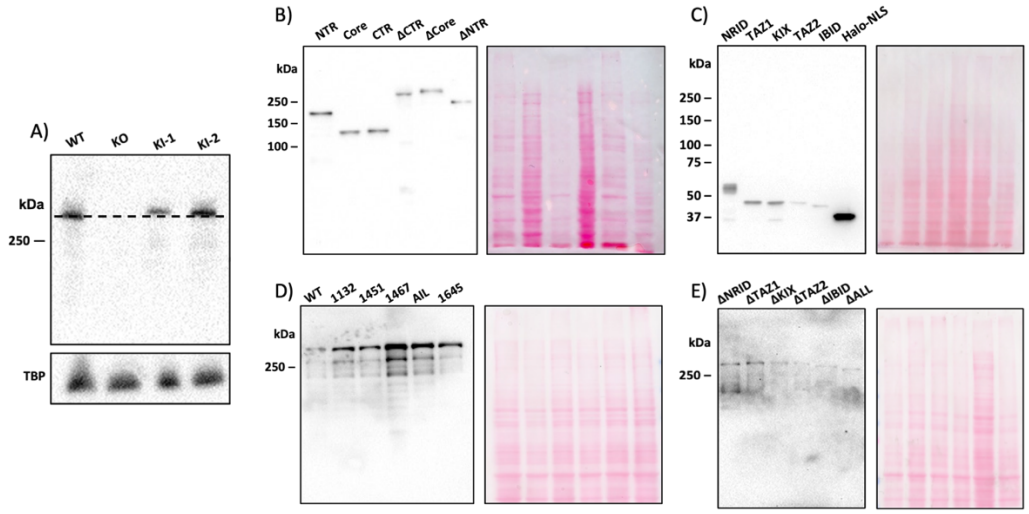

Figure 3.S1—Western blots of p300 variants.
(A) Anti-p300 blot of unedited U2OS cells, the *EP300* knockout line, and the two Halo-p300 knock-in lines, with loading control beneath. Dashed line is to show shift in molecular weight from addition of HaloTag. (B-E) Anti-FLAG blots of all transgene constructs in the paper with Ponceau stains to the right as loading controls. Constructs were loaded to achieve comparable intensities—for relative expression levels, see Fig. 3.S2.

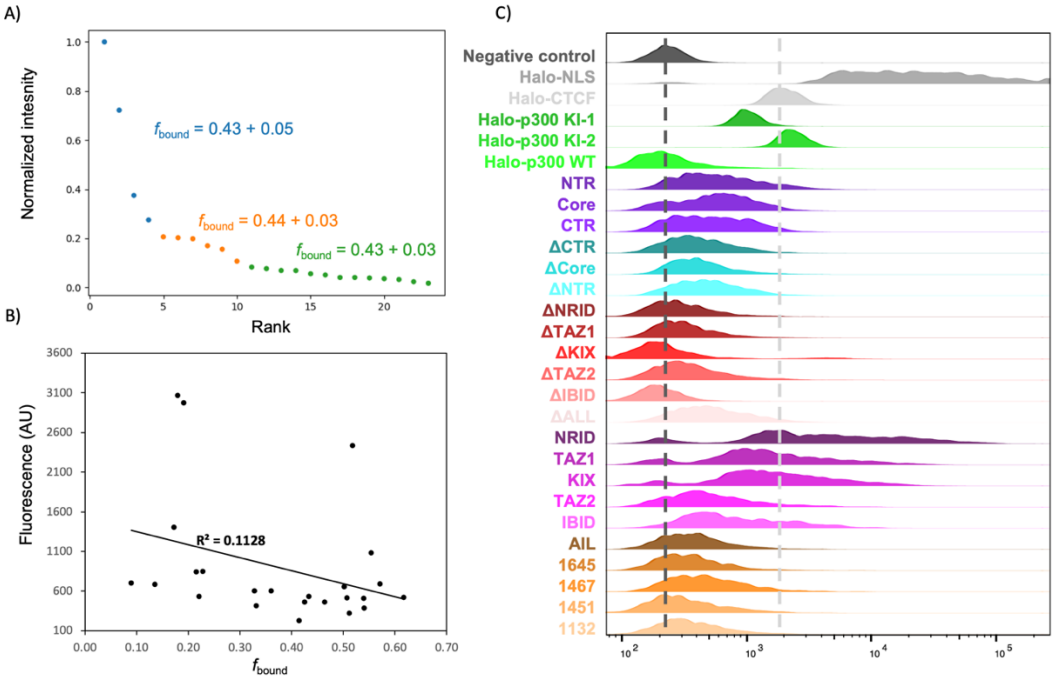

Figure 3.S2—Fraction bound is unaffected by p300 expression level.
(A) Plot of cellular intensities from a transient transfection of WT p300 into *EP300*-knockout cells with SMT results of three intensity bins overlaid. (B) All p300 mutants in the paper were measured by flow cytometry (C) and their mean cellular fluorescence plotted against SMT-derived f_{bound}.

p300 Core is neither necessary nor sufficient for chromatin binding

To address which domains mediate p300 chromatin engagement, we first assessed which of its three regions (NTR, Core, CTR) is sufficient for binding (Figure 3.2B). Strikingly, the Core had essentially no ability to bind on its own (f_{bound} = 0.09 ± 0.01). The NTR exhibited a modest capacity to bind (f_{bound} = 0.22 ± 0.02), while CTR was sufficient to reach full-length p300 levels of binding (f_{bound} = 0.50 ± 0.02). Both NTR and Core constructs had significantly reduced residence times (t = 16, 14 s) compared to CTR and p300 (t = 26, 26 s) (Figure 3.S3).

We then asked which of the domains is necessary for binding by tracking truncations of p300: NTR-Core (ΔCTR), Core-CTR (ΔNTR), and NTR-CTR (ΔCore) (Figure 3.2D). Remarkably, ΔCore had a somewhat enhanced f_{bound} of 0.62 ± 0.03—a surprising result we return to later. ΔNTR and ΔCTR showed significantly decreased ability to bind chromatin (f_{bound} = 0.36 ± 0.03, 0.22 ± 0.02) compared to WT p300 (fbound = 0.51 ± 0.02). Together, these results suggest that the ChID-containing Core is dispensable for p300–chromatin binding in vivo while the NTR and CTR are necessary and sufficient (albeit to different extents). To confirm this, we performed SMT on a series of Core mutants which affect p300 binding and activity in vitro (including a bromodomain mutant) and saw no substantial changes in our in vivo assay (Figure 3.S4).

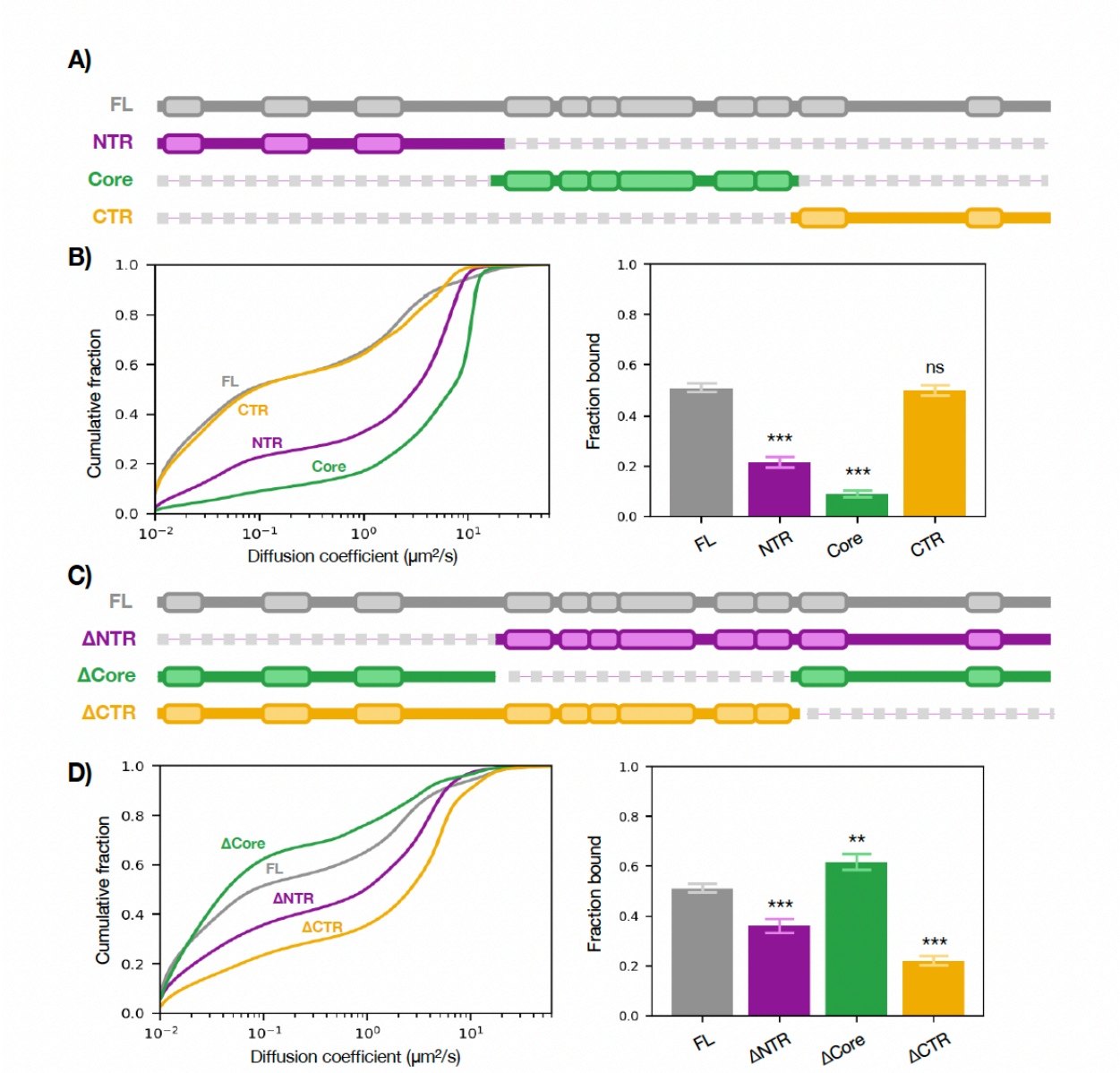

Figure 3.2—Sufficiency and necessity of p300 regions for chromatin binding.
(A) Schematic of p300 regions. (B) SMT of isolated regions of p300 (left) with summary bar plot of f_{bound} (right). (C) Schematic of p300 truncations. (D) SMT of truncations of p300 (left) with summary bar plot of f_{bound} (right). Bars represent bootstrapping mean ± SD while p-values are reported as ns, (not significant), * (< 0.05), ** (<0.01), or *** (<0.001). Note that the full-length "FL" data are those identified as "WT Transgene" in Figure 1. See also Fig. 3.S3.

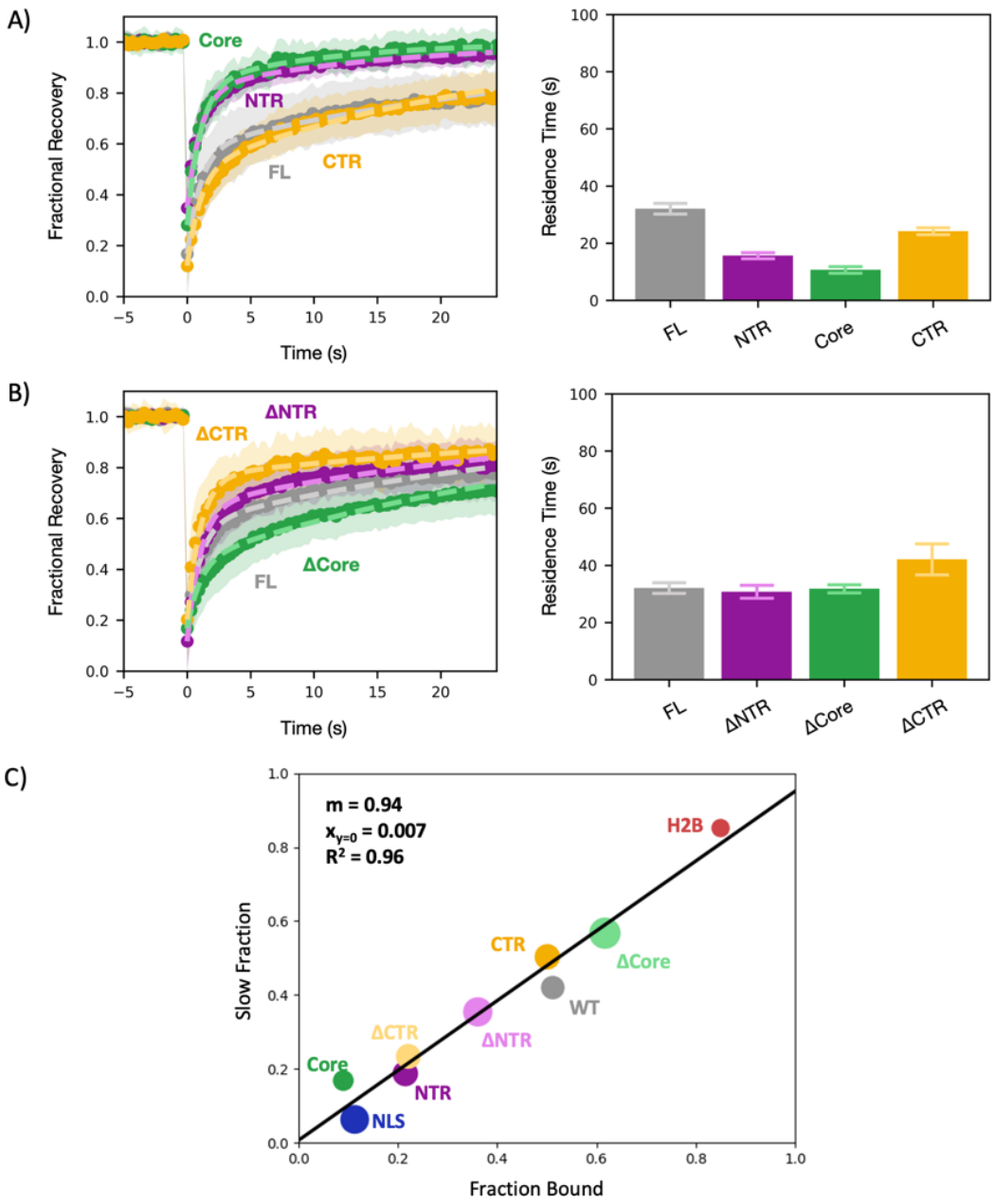

Figure 3.S3—FRAP confirms SMT findings.
(A) FRAP plots of p300 regions (left) with residence times plotted (right). (B) FRAP plots of p300 truncations (left) with residence times plotted (right). (C) Scatter plot of fraction bound (SMT) and slow fraction (FRAP) for transgene constructs shows a high degree of correlation. The size of each point encompasses both the standard deviation of the bound fraction and the 95% confidence interval of the slow fraction.

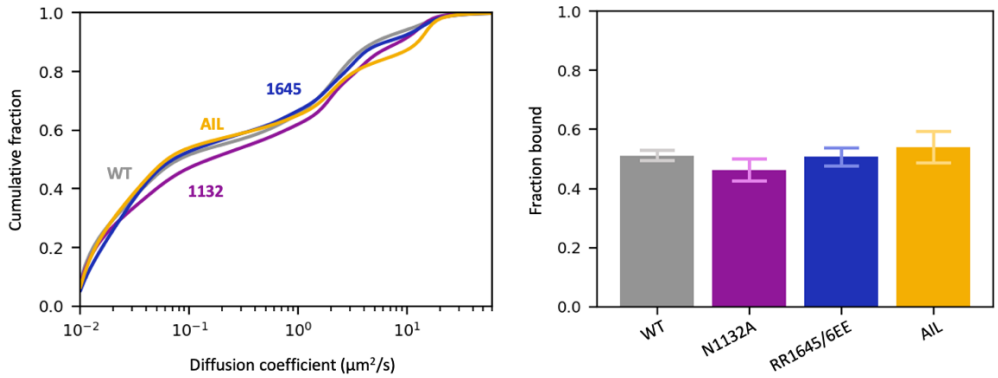

Figure 3.S4—SMT of Core mutants.
SMT plots as in Figure 1 of Core mutants shown in vitro to have large effects on p300 catalytic activity: N1132A, RR1645/6EE, autoinhibitory loop $[K>R]_9$ (AIL).

p300 TFIDs are necessary for chromatin binding

The finding that p300's chromatin-binding capacity lies outside its Core domain does not necessarily mean that the TFIDs are what mediate binding as it could be the IDRs, which compose the vast majority of the NTR and CTR. (Whereas each TFID is 50-80 aa, there is approximately 1,400 aa of IDR.) We therefore measured the sufficiency of the TFIDs for chromatin binding by expressing them as HaloTag fusions and observed that each TFID has only a modest capacity to bind when acting alone (Figure 3.3B). Next, we tested the necessity of these domains by deleting each of the TFIDs as well as all the TFIDs in the otherwise full-length protein and performed SMT (Figure 3.3D). Although each TFID deletion only partially impaired chromatin binding, when all TFIDs were deleted (ΔALL), there was a drastic reduction in f_{bound} to just above the baseline (0.14 ± 0.02). Hence, it appears that the combined action of multiple TFIDs is required to bring p300 to chromatin. That deleting all TFIDs essentially incapacitated chromatin binding also suggests that the IDRs are not sufficient for nor a major contributor to p300–chromatin association. Additionally, the low f_{bound} of the ΔALL construct is another indication that the Core is not sufficient for p300-chromatin binding.

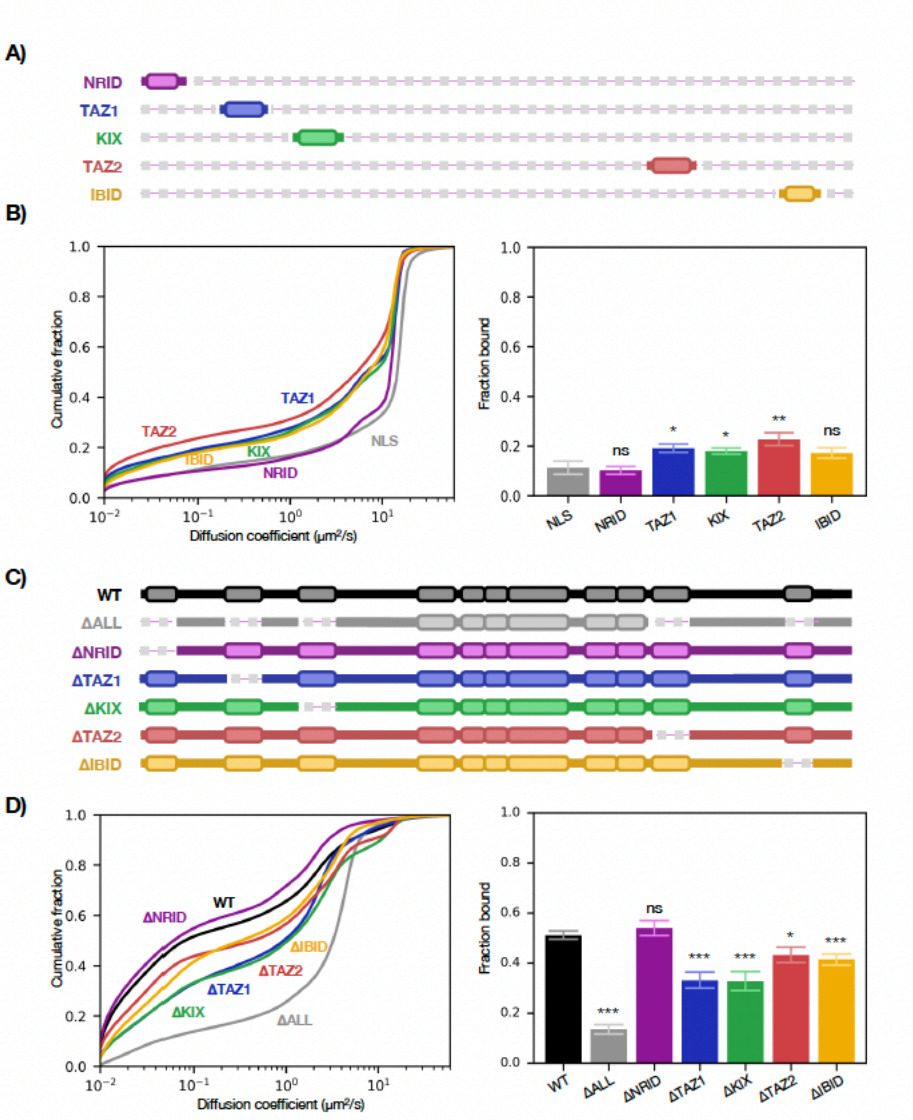

Figure 3.3—Sufficiency and necessity of p300 TFIDs for chromatin binding.
A) Schematic of p300 TF interaction domains (core region in grey). (B) SMT of TFIDs (left) with summary bar plot of f_{bound} (right). (C) Schematic of p300 TFID deletions. (D) SMT of TFID deletions of p300 (left) with summary bar plot of f_{bound} (right). Bars represent bootstrapping mean ± SD while p-values are reported as ns, (not significant), * (< 0.05), ** (<0.01), or *** (<0.001). Note that the full-length "FL" data are those identified as "WT Transgene" in Figure 3.1.

Acetyltransferase activity opposes p300–chromatin binding

Although it seems that the Core does not contribute appreciably to p300–chromatin binding, it is intriguing that its deletion increased p300 f_{bound} from 0.51 ± 0.02 to 0.62 ± 0.03. We wondered whether this was the consequence of the loss of its acetyltransferase activity and tested this in two ways: by imaging a catalytically dead mutant (Y1467F)[141] and tracking WT p300 after addition of the potent and specific catalytic inhibitor A485.[142] Both slightly increased p300 f_{bound} similar to complete loss of the Core domain (Figure 3.4A), although neither change reached statistical significance. A485 treatment likewise led to slower FRAP recovery, supporting an increase in

chromatin binding upon loss of catalytic activity. To test whether this was a direct effect of p300 activity and not an indirect consequence of some off-target cellular response, we designed a p300 point mutant that remains catalytically active in the presence of A485 (H1451K) (manuscript in preparation) and as expected, saw no change by FRAP (Figure 3.4B) or SMT (Figure 3.S5) upon A485 addition. These results tally with previous ChIP-seq data showing an increase of p300 peak heights upon A485 treatment.[100] and prior in vitro data suggesting that p300 dissociation from Mediator is dependent on p300 acetyltransferase activity.[143]

The finding that catalytic activity opposes chromatin binding provided a possible explanation for two curious observations: ΔNTR (Core-CTR) had reduced binding compared to CTR and deletion of the two N-terminal TFIDs had a greater impact on f_{bound} than deletion of the C-terminal TFIDs—both of which are strange because the CTR is both necessary and sufficient for full chromatin association. We therefore reasoned that the N-terminus may function (directly or indirectly) to inhibit the effect of Core catalytic activity on CTR binding. Indeed, the reduction in f_{bound} from adding the Core to the CTR was rescued by treatment with A485 (Figure 3.4C). Furthermore, ΔTAZ1 and ΔKIX were both hyper-sensitive to the drug compared to the FL with respect to f_{bound} (Figure 3.4D), indicating that the N-terminal TFIDs counteract the effect of Core catalytic activity on p300-chromatin interaction.

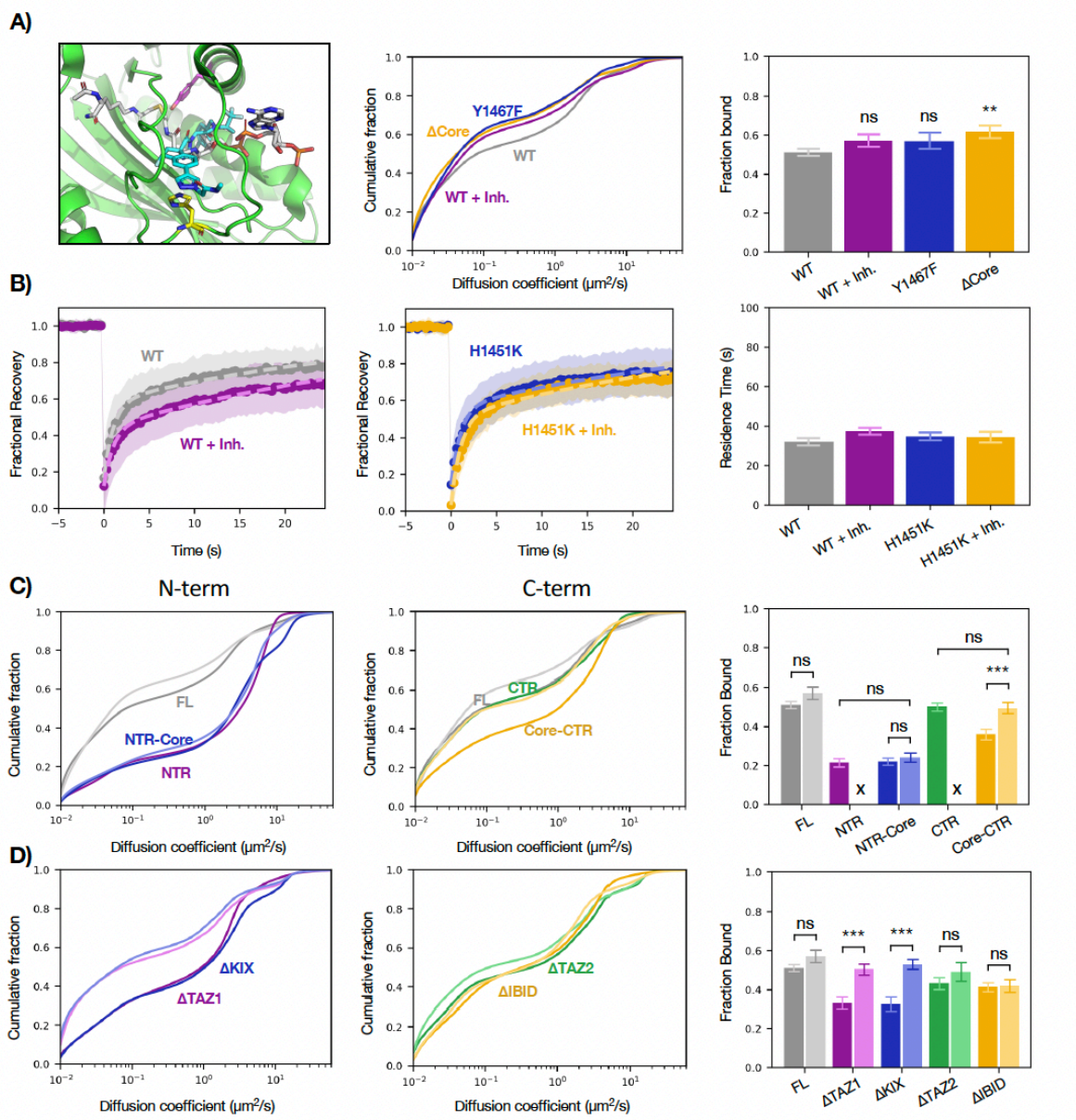

Figure 3.4—Effects of the catalytic core on p300–chromatin interactions.
(A) Cartoon representation (left) of p300 active site showing residues H1451 (pink) and Y1467 (yellow), substrate mimetic (grey), and A485 inhibitor (teal). Plots comparing diffusion spectra (center) of WT p300 and three perturbations of p300 catalytic activity: addition of inhibitor A485, active site mutation Y1467F, and deletion of the catalytic core, with bar plot of f_{bound} (right). (B) FRAP plots of WT (left) and mutant (center) in response to A485 with residence times plotted for each construct (right). Note that in both A and B, the "WT" data are those identified as "WT Transgene" in Figure 3.1. (C) SMT of NTR-containing (left) and CTR-containing (center) constructs in the presence (lighter hue) or absence (darker hue) of A485, with bar plot of f_{bound} (right). [X] Data not acquired. (D) SMT of N-terminal ΔTFID (left) and C-terminal ΔTFID (center) constructs in the presence (lighter hue) or absence (darker hue) of A485, with bar plot of f_{bound} (right). Note that in both C and D, the "FL" data are those identified as "WT Transgene" in Figure 1 and "WT" in A and B. SMT: bars represent bootstrapping mean ± SD while p-values are reported as ns, (not significant), * (< 0.05), ** (<0.01), or *** (<0.001). FRAP: Bars represent best fit ± 95% CI. See also Fig. 3.S5.

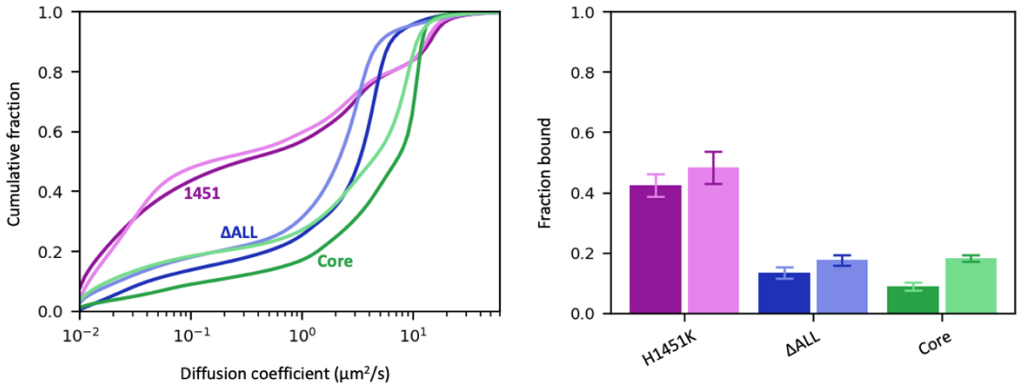

Figure 3.S5—SMT of Core mutants.
SMT plots of mutants ± A485: the inhibitor-resistant mutant 1451 has no significant change, while the two unbound constructs ΔALL and Core show mild increases upon addition of the inhibitor.

Discussion

Combining high-speed SMT with the domain-mapping strategy classically used in in vitro biochemistry allowed us to determine which domains of a modular coactivator mediate its chromatin binding in living cells. The results suggest that p300 binds chromatin primarily through multivalent TF–TFID interactions, not by its ChIDs or IDRs (Figure 3.5A).

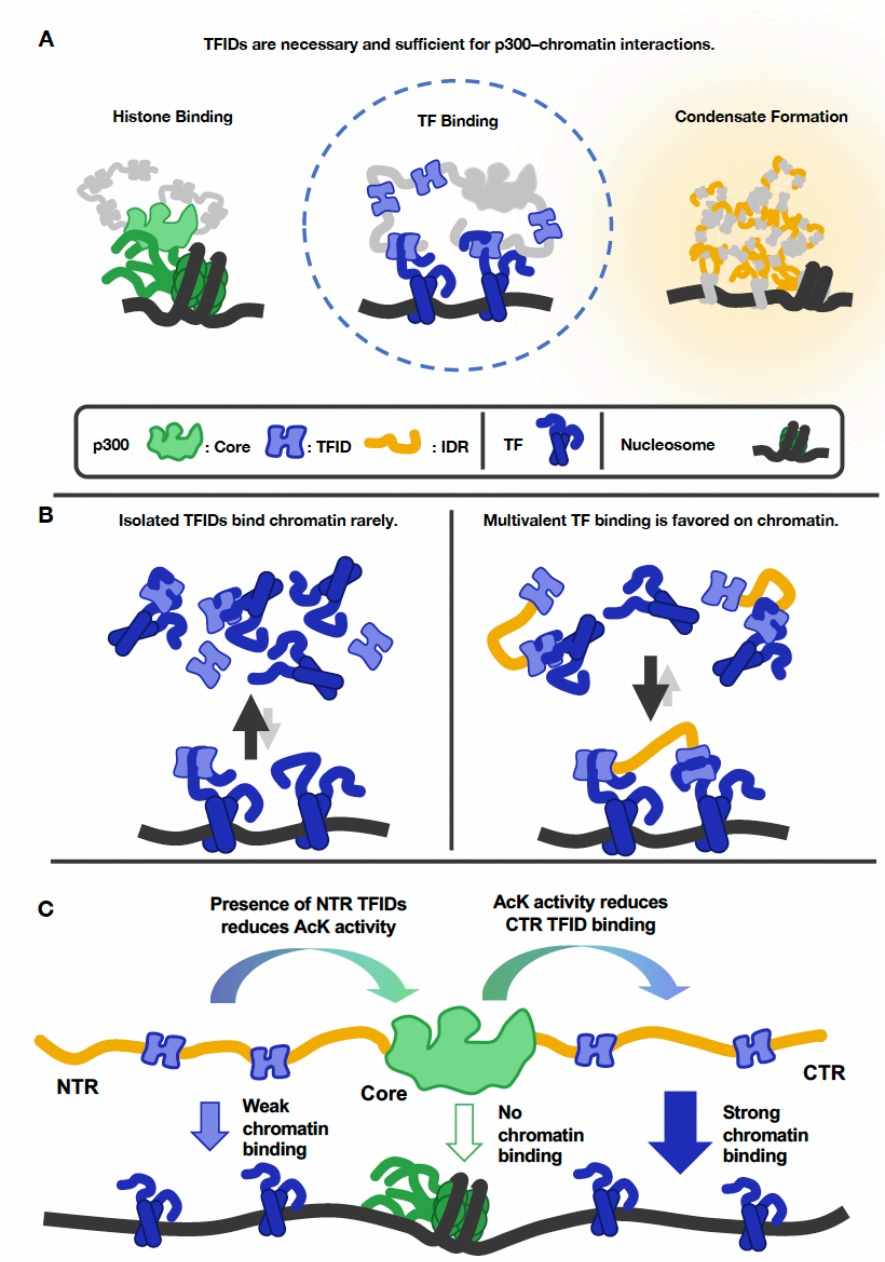

Figure 3.5— Graphical model of p300–chromatin interactions.
(A) Three existing models of how p300 engages with chromatin, the second of which is strongly supported by in vivo SMT data. (B) Multivalent TFID-TF interactions favor the chromatin-bound state. (C) The contributions of each p300 domain to p300–chromatin interactions.

SMT discriminates between models of p300–chromatin engagement

Two measurements we made are particularly clarifying: Deleting the Core increased f_{bound} to 0.62, while removal of the five small TFIDs essentially eliminated p300 chromatin binding, bringing f_{bound} down to 0.14 (Figure 3.3). Thus, the collective action of the small helical peptides we call TFIDs is both necessary and sufficient for p300–chromatin interactions. While the Core and the IDRs undoubtedly contribute to other functions, our data consistently show that the alleged binding of the Core to histones is neither necessary nor sufficient and that purported IDR-mediated interactions are not sufficient for nor a major driver of p300's interaction with chromatin. Instead, we observe that TFID-mediated multivalent interactions with sequence-specific TFs are necessary and sufficient to support p300-chromatin interaction, in line with previous in vitro findings.[129,144] (Though it is possible, of course, that factors other than sequence-specific TFs also engage p300 TFIDs.) These findings comport with in vitro binding assays showing that bromodomains have micromolar affinity for acetylated histone peptides[137] while TFIDs have nanomolar affinity for TF peptides.[145,146] Furthermore, the increased f_{bound} of ΔCore relative to WT suggests that catalytic activity does not cause p300–chromatin association through its bromodomain, which is confirmed by the finding that chromatin binding ability was not impaired by the acetyl-transferase inhibitor A485, a catalytically dead mutant, or the p300 bromodomain mutant (Figure 3.4B, Figure 3.S4).

Multiple TFID interactions enable avid-like binding of p300 to chromatin

Beyond addressing our initial questions, our investigation yielded insights into the mechanism of p300 recruitment. The finding that each TFID on its own largely occupies the unbound state suggests that the majority of each individual TFID's binding partners are unbound, and that a single TFID has little or no preference for chromatin-bound TFs. (We consider it unlikely that the TFID is merely unbound by a TF, given the multiplicity of TF binding partners and the aforementioned nanomolar affinities measured between TFs and p300 TFIDs.) Multimerization of TFIDs can strongly favor interactions with chromatin, as demonstrated by the fact that a two-TFID fragment, the CTR, has a dramatically greater chromatin-bound fraction than either C-terminal TFID alone. A simple explanation for this is an avidity effect of engaging multiple TFs bound to adjacent DNA sites. Binding multiple TFs on chromatin is energetically more favorable than binding them in solution, as the entropic cost of bringing the TFs together is already paid. In principle, a coactivator with multiple TFIDs could even have a higher chromatin-bound fraction than any of the TFs it binds.

p300 is an obligate integrator of combinatorial TF inputs

The requisite for multiple TFIDs in p300 recruitment points to the important finding that p300 functions as an obligate integrator—that is, it not only *can* but *must* bind multiple TFs, or a single TF with high valency, to associate with chromatin. Such integration has profound implications for p300 recruitment at two levels—in trans and in cis. In trans, any gene-regulatory process that requires the action of p300 must deploy a TF or multiple TFs that can engage multiple TFIDs. In cis, any regulatory element that requires p300 activity must be able to bind more than one of its TF partners, rendering isolated TF binding sequences insufficient for recruiting p300 in a competitive environment. If this is a general rule for coactivator recruitment, the high degeneracy of eukaryotic TF recognition motifs[147] might be straightforwardly rationalized: Because a TF must act through recruiting coactivators in coordination with another TF, their modest individual specificities are multiplied to recruit coactivators to cis regulatory elements that are capable of

binding both TFs simultaneously—thus solving the problem we posed of how a limiting pool of coactivators is distributed among target loci. Those factors that can engage more TFIDs on their own, such as p53,[148] may be the more context-independent activators, while those that can only engage one of the TFIDs are likely more dependent on other proteins (i.e., additional TFs) to achieve activation of target genes.

Last, because p300 binding is sensitive to catalytic activity, our assay was able to uncover p300 catalytic regulation by its N-terminal TFIDs (Figure 3.4C-D). The Core is most active in the absence of TAZ1 and KIX, which suggests that these domains or their binding partners either directly inhibit acetyltransferase activity or modulate its effect on chromatin binding of the C-terminal TFIDs (Figure 3.5C). While these data do not shed light on the mechanism of such regulation, there is an important biological consequence of the observation: That p300 catalytic activity is regulable through two of its TFIDs shows another way in which it is an obligate integrator of TF inputs.

References

1. Nicholson, D.J. (2019). Is the cell really a machine? J Theor Biol *477*, 108–126. 10.1016/j.jtbi.2019.06.002.
2. Garcia, D.A., Fettweis, G., Presman, D.M., Paakinaho, V., Jarzynski, C., Upadhyaya, A., and Hager, G.L. (2021). Power-law behavior of transcription factor dynamics at the single-molecule level implies a continuum affinity model. Nucleic Acids Res *49*, 6605–6620. 10.1093/nar/gkab072.
3. Darzacq, X., Yao, J., Larson, D.R., Causse, S.Z., Bosanac, L., de Turris, V., Ruda, V.M., Lionnet, T., Zenklusen, D., Guglielmi, B., et al. (2009). Imaging Transcription in Living Cells. Annu Rev Biophys *38*, 173–196. 10.1146/annurev.biophys.050708.133728.
4. Hager, G.L., McNally, J.G., and Misteli, T. (2009). Transcription Dynamics. Mol Cell *35*, 741–753. 10.1016/j.molcel.2009.09.005.
5. Neumayr, C., Haberle, V., Serebreni, L., Karner, K., Hendy, O., Boija, A., Henninger, J.E., Li, C.H., Stejskal, K., Lin, G., et al. (2022). Differential cofactor dependencies define distinct types of human enhancers. Nature *606*, 406–413. 10.1038/s41586-022-04779-x.
6. Demmerle, J., Hao, S., and Cai, D. (2023). Transcriptional condensates and phase separation: condensing information across scales and mechanisms. Nucleus *14*. 10.1080/19491034.2023.2213551.
7. Xie, L., Dong, P., Qi, Y., Hsieh, T.-H.S., English, B.P., Jung, S., Chen, X., De Marzio, M., Casellas, R., Chang, H.Y., et al. (2022). BRD2 compartmentalizes the accessible genome. Nat Genet *54*, 481–491. 10.1038/s41588-022-01044-9.
8. Ferrie, J.J., Karr, J.P., Tjian, R., and Darzacq, X. (2022). "Structure"-function relationships in eukaryotic transcription factors: The role of intrinsically disordered regions in gene regulation. Mol Cell *82*, 3970–3984. 10.1016/j.molcel.2022.09.021.
9. Turing, A.M. (1952). The chemical basis of morphogenesis. Philos Trans R Soc Lond B Biol Sci *237*, 37–72. 10.1098/rstb.1952.0012.
10. Kerszberg, M., and Wolpert, L. (1998). Mechanisms for Positional Signalling by Morphogen Transport: a Theoretical Study. J Theor Biol *191*, 103–114. 10.1006/jtbi.1997.0575.
11. Tunnacliffe, E., and Chubb, J.R. (2020). What Is a Transcriptional Burst? Trends in Genetics *36*, 288–297. 10.1016/j.tig.2020.01.003.
12. Huang, S. (2022). Towards a unification of the 2 meanings of "epigenetics." PLoS Biol *20*, e3001944. 10.1371/journal.pbio.3001944.
13. Benveniste, D., Sonntag, H.-J., Sanguinetti, G., and Sproul, D. (2014). Transcription factor binding predicts histone modifications in human cell lines. Proceedings of the National Academy of Sciences *111*, 13367–13372. 10.1073/pnas.1412081111.
14. Ho, S.N., Biggar, S.R., Spencer, D.M., Schreiber, S.L., and Crabtree, G.R. (1996). Dimeric ligands define a role for transcriptional activation domains in reinitiation. Nature *382*, 822–826. 10.1038/382822a0.
15. Hansen, J.L., Loell, K.J., and Cohen, B.A. (2022). A test of the pioneer factor hypothesis using ectopic liver gene activation. Elife *11*. 10.7554/eLife.73358.

16. Mirny, L.A. (2010). Nucleosome-mediated cooperativity between transcription factors. Proceedings of the National Academy of Sciences *107*, 22534–22539. 10.1073/pnas.0913805107.
17. Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676. 10.1016/j.cell.2006.07.024.
18. Bluestone, J.A., Mackay, C.R., O'Shea, J.J., and Stockinger, B. (2009). The functional plasticity of T cell subsets. Nat Rev Immunol *9*, 811–816. 10.1038/nri2654.
19. Murphy, K.M., and Stockinger, B. (2010). Effector T cell plasticity: flexibility in the face of changing circumstances. Nat Immunol *11*, 674–680. 10.1038/ni.1899.
20. Gupta, P.B., Pastushenko, I., Skibinski, A., Blanpain, C., and Kuperwasser, C. (2019). Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. Cell Stem Cell *24*, 65–78. 10.1016/j.stem.2018.11.011.
21. Black, J.R.M., and McGranahan, N. (2021). Genetic and non-genetic clonal diversity in cancer evolution. Nat Rev Cancer *21*, 379–392. 10.1038/s41568-021-00336-2.
22. Househam, J., Heide, T., Cresswell, G.D., Spiteri, I., Kimberley, C., Zapata, L., Lynn, C., James, C., Mossner, M., Fernandez-Mateos, J., et al. (2022). Phenotypic plasticity and genetic control in colorectal cancer evolution. Nature *611*, 744–753. 10.1038/s41586-022-05311-x.
23. Whiting, F.J.H., Househam, J., Baker, A.-M., Sottoriva, A., and Graham, T.A. (2023). Phenotypic noise and plasticity in cancer evolution. Trends Cell Biol. 10.1016/j.tcb.2023.10.002.
24. Burkhardt, D.B., San Juan, B.P., Lock, J.G., Krishnaswamy, S., and Chaffer, C.L. (2022). Mapping Phenotypic Plasticity upon the Cancer Cell State Landscape Using Manifold Learning. Cancer Discov *12*, 1847–1859. 10.1158/2159-8290.CD-21-0282.
25. Rosenfeld, M.G. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. Genes Dev *20*, 1405–1428. 10.1101/gad.1424806.
26. de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev *26*, 11–24. 10.1101/gad.179804.111.
27. Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. Mol Cell *78*, 539-553.e8. 10.1016/j.molcel.2020.03.002.
28. Hoffman, D.P., Shtengel, G., Xu, C.S., Campbell, K.R., Freeman, M., Wang, L., Milkie, D.E., Pasolli, H.A., Iyer, N., Bogovic, J.A., et al. (2020). Correlative three-dimensional super-resolution and block-face electron microscopy of whole vitreously frozen cells. Science (1979) *367*. 10.1126/science.aaz5357.
29. Teves, S.S., An, L., Hansen, A.S., Xie, L., Darzacq, X., and Tjian, R. (2016). A dynamic mode of mitotic bookmarking by transcription factors. Elife *5*. 10.7554/eLife.22280.
30. Liu, Z., and Tjian, R. (2018). Visualizing transcription factor dynamics in living cells. Journal of Cell Biology *217*, 1181–1191. 10.1083/jcb.201710038.

31. Brandão, H.B., Gabriele, M., and Hansen, A.S. (2021). Tracking and interpreting long-range chromatin interactions with super-resolution live-cell imaging. Curr Opin Cell Biol *70*, 18–26. 10.1016/j.ceb.2020.11.002.
32. Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nat Genet *50*, 1296–1303. 10.1038/s41588-018-0175-z.
33. Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. Mol Cell *76*, 473-484.e7. 10.1016/j.molcel.2019.07.038.
34. Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. Elife *8*. 10.7554/eLife.41769.
35. Mateo, L.J., Murphy, S.E., Hafner, A., Cinquini, I.S., Walker, C.A., and Boettiger, A.N. (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. Nature *568*, 49–54. 10.1038/s41586-019-1035-4.
36. Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944.e22. 10.1016/j.cell.2017.05.004.
37. Luan, J., Xiang, G., Gómez-García, P.A., Tome, J.M., Zhang, Z., Vermunt, M.W., Zhang, H., Huang, A., Keller, C.A., Giardine, B.M., et al. (2021). Distinct properties and functions of CTCF revealed by a rapidly inducible degron system. Cell Rep *34*, 108783. 10.1016/j.celrep.2021.108783.
38. Rao, S.S.P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320.e24. 10.1016/j.cell.2017.09.026.
39. Hansen, A.S., Hsieh, T.-H.S., Cattoglio, C., Pustova, I., Saldaña-Meyer, R., Reinberg, D., Darzacq, X., and Tjian, R. (2019). Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. Mol Cell *76*, 395-411.e13. 10.1016/j.molcel.2019.07.039.
40. Xiao, J.Y., Hafner, A., and Boettiger, A.N. (2021). How subtle changes in 3D structure can create large changes in transcription. Elife *10*. 10.7554/eLife.64320.
41. Biggar, K.K., and Li, S.S.-C. (2015). Non-histone protein methylation as a regulator of cellular signalling and function. Nat Rev Mol Cell Biol *16*, 5–17. 10.1038/nrm3915.
42. Narita, T., Weinert, B.T., and Choudhary, C. (2019). Functions and mechanisms of non-histone protein acetylation. Nat Rev Mol Cell Biol *20*, 156–174. 10.1038/s41580-018-0081-3.
43. Langeberg, L.K., and Scott, J.D. (2015). Signalling scaffolds and local organization of cellular behaviour. Nat Rev Mol Cell Biol *16*, 232–244. 10.1038/nrm3966.
44. Karlic, R., Chung, H.-R., Lasserre, J., Vlahovicek, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proceedings of the National Academy of Sciences *107*, 2926–2931. 10.1073/pnas.0909344107.

45. Fulco, C.P., Nasser, J., Jones, T.R., Munson, G., Bergman, D.T., Subramanian, V., Grossman, S.R., Anyoha, R., Doughty, B.R., Patwardhan, T.A., et al. (2019). Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. Nat Genet *51*, 1664–1669. 10.1038/s41588-019-0538-0.
46. Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol *33*, 510–517. 10.1038/nbt.3199.
47. Raisner, R., Kharbanda, S., Jin, L., Jeng, E., Chan, E., Merchant, M., Haverty, P.M., Bainer, R., Cheung, T., Arnott, D., et al. (2018). Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation. Cell Rep *24*, 1722–1729. 10.1016/j.celrep.2018.07.041.
48. Zhang, T., Zhang, Z., Dong, Q., Xiong, J., and Zhu, B. (2020). Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. Genome Biol *21*, 45. 10.1186/s13059-020-01957-w.
49. Catarino, R.R., and Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. Genes Dev *32*, 202–223. 10.1101/gad.310367.117.
50. Dancy, B.M., and Cole, P.A. (2015). Protein Lysine Acetylation by p300/CBP. Chem Rev *115*, 2419–2452. 10.1021/cr500452k.
51. Weinert, B.T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B.K., Schölz, C., Hamilton, W.B., Zucconi, B.E., Wang, W.W., Liu, W.R., et al. (2018). Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. Cell *174*, 231-244.e12. 10.1016/j.cell.2018.04.033.
52. Glozak, M.A., Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. Gene *363*. 10.1016/j.gene.2005.09.010.
53. Sivanand, S., Viney, I., and Wellen, K.E. (2018). Spatiotemporal Control of Acetyl-CoA Metabolism in Chromatin Regulation. Trends Biochem Sci *43*, 61–74. 10.1016/j.tibs.2017.11.004.
54. Izeddin, I., Récamier, V., Bosanac, L., Cissé, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Bénichou, O., Voituriez, R., Bensaude, O., et al. (2014). Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. Elife *3*. 10.7554/eLife.02230.
55. Bénichou, O., Chevalier, C., Klafter, J., Meyer, B., and Voituriez, R. (2010). Geometry-controlled kinetics. Nat Chem *2*, 472–477. 10.1038/nchem.622.
56. Chen, J., Zhang, Z., Li, L., Chen, B.-C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., et al. (2014). Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. Cell *156*, 1274–1285. 10.1016/j.cell.2014.01.062.
57. Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science (1979) *361*, 412–415. 10.1126/science.aar4199.
58. Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-

complexity domain interactions that control gene transcription. Science (1979) *361*, eaar2555. 10.1126/science.aar2555.
59. Liu, Z., Legant, W.R., Chen, B.-C., Li, L., Grimm, J.B., Lavis, L.D., Betzig, E., and Tjian, R. (2014). 3D imaging of Sox2 enhancer clusters in embryonic stem cells. Elife *3*. 10.7554/eLife.04236.
60. Mir, M., Reimer, A., Haines, J.E., Li, X.-Y., Stadler, M., Garcia, H., Eisen, M.B., and Darzacq, X. (2017). Dense Bicoid hubs accentuate binding along the morphogen gradient. Genes Dev *31*, 1784–1794. 10.1101/gad.305078.117.
61. Tsai, A., Muthusamy, A.K., Alves, M.R., Lavis, L.D., Singer, R.H., Stern, D.L., and Crocker, J. (2017). Nuclear microenvironments modulate transcription from low-affinity enhancers. Elife *6*. 10.7554/eLife.28975.
62. Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell *166*, 358–368. 10.1016/j.cell.2016.05.025.
63. Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., and Spitz, F. (2016). The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. Dev Cell *39*, 529–543. 10.1016/j.devcel.2016.10.015.
64. Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854–858. 10.1038/nature07730.
65. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311–318. 10.1038/ng1966.
66. Giordano, A., and Avantaggiati, M.L. (1999). p300 and CBP: Partners for life and death. J Cell Physiol *181*, 218–230. 10.1002/(SICI)1097-4652(199911)181:2<218::AID-JCP4>3.0.CO;2-5.
67. Gillespie, M.A., Palii, C.G., Sanchez-Taltavull, D., Shannon, P., Longabaugh, W.J.R., Downes, D.J., Sivaraman, K., Espinoza, H.M., Hughes, J.R., Price, N.D., et al. (2020). Absolute Quantification of Transcription Factors Reveals Principles of Gene Regulation in Erythropoiesis. Mol Cell *78*, 960-974.e11. 10.1016/j.molcel.2020.03.031.
68. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences *107*, 21931–21936. 10.1073/pnas.1016071107.
69. Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283. 10.1038/nature09692.
70. Thompson, P.R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R.G., Wong, J., et al. (2004). Regulation of the p300 HAT domain via a novel activation loop. Nat Struct Mol Biol *11*, 308–315. 10.1038/nsmb740.
71. Ortega, E., Rengachari, S., Ibrahim, Z., Hoghoughi, N., Gaucher, J., Holehouse, A.S., Khochbin, S., and Panne, D. (2018). Transcription factor dimerization activates the p300 acetyltransferase. Nature *562*, 538–544. 10.1038/s41586-018-0621-1.

72. Whitmarsh, A.J., and Davis, R.J. (2000). Regulation of transcription factor function by phosphorylation. Cellular and Molecular Life Sciences *57*, 1172–1183. 10.1007/PL00000757.
73. Kaypee, S., Sahadevan, S.A., Patil, S., Ghosh, P., Roy, N.S., Roy, S., and Kundu, T.K. (2018). Mutant and Wild-Type Tumor Suppressor p53 Induces p300 Autoacetylation. iScience *4*, 260–272. 10.1016/j.isci.2018.06.002.
74. Mayr, B., and Montminy, M. (2001). Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol *2*, 599–609. 10.1038/35085068.
75. Zhong, H., May, M.J., Jimi, E., and Ghosh, S. (2002). The Phosphorylation Status of Nuclear NF-KB Determines Its Association with CBP/p300 or HDAC-1. Mol Cell *9*, 625–636. 10.1016/S1097-2765(02)00477-X.
76. Wathelet, M.G., Lin, C.H., Parekh, B.S., Ronco, L. V, Howley, P.M., and Maniatis, T. (1998). Virus Infection Induces the Assembly of Coordinately Activated Transcription Factors on the IFN-β Enhancer In Vivo. Mol Cell *1*, 507–518. 10.1016/S1097-2765(00)80051-9.
77. Schuringa, J.-J., Schepers, H., Vellenga, E., and Kruijer, W. (2001). Ser727-dependent transcriptional activation by association of p300 with STAT3 upon IL-6 stimulation. FEBS Lett *495*, 71–76. 10.1016/S0014-5793(01)02354-7.
78. Major, M.L., Lepe, R., and Costa, R.H. (2004). Forkhead Box M1B Transcriptional Activity Requires Binding of Cdk-Cyclin Complexes for Phosphorylation-Dependent Recruitment of p300/CBP Coactivators. Mol Cell Biol *24*, 2649–2661. 10.1128/MCB.24.7.2649-2661.2004.
79. Simonsson, M., Kanduri, M., Grönroos, E., Heldin, C.-H., and Ericsson, J. (2006). The DNA Binding Activities of Smad2 and Smad3 Are Regulated by Coactivator-mediated Acetylation. Journal of Biological Chemistry *281*, 39870–39880. 10.1074/jbc.M607868200.
80. Freedman, S.J., Sun, Z.-Y.J., Poy, F., Kung, A.L., Livingston, D.M., Wagner, G., and Eck, M.J. (2002). Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1. Proceedings of the National Academy of Sciences *99*, 5367–5372. 10.1073/pnas.082117899.
81. Vahedi, G., Takahashi, H., Nakayamada, S., Sun, H., Sartorelli, V., Kanno, Y., and O'Shea, J.J. (2012). STATs Shape the Active Enhancer Landscape of T Cell Populations. Cell *151*, 981–993. 10.1016/j.cell.2012.09.044.
82. Khorasanizadeh, S., and Rastinejad, F. (2001). Nuclear-receptor interactions on DNA-response elements. Trends Biochem Sci *26*, 384–390. 10.1016/S0968-0004(01)01800-X.
83. Dornan, D., Shimizu, H., Perkins, N.D., and Hupp, T.R. (2003). DNA-dependent Acetylation of p53 by the Transcription Coactivator p300. Journal of Biological Chemistry *278*, 13431–13441. 10.1074/jbc.M211460200.
84. Češková, P., Chichger, H., Wallace, M., Vojtesek, B., and Hupp, T.R. (2006). On the Mechanism of Sequence-specific DNA-dependent Acetylation of p53: The Acetylation Motif is Exposed upon DNA Binding. J Mol Biol *357*, 442–456. 10.1016/j.jmb.2005.12.026.
85. Zaware, N., and Zhou, M.-M. (2019). Bromodomain biology and drug discovery. Nat Struct Mol Biol *26*, 870–879. 10.1038/s41594-019-0309-8.

86. Lamonica, J.M., Deng, W., Kadauke, S., Campbell, A.E., Gamsjaeger, R., Wang, H., Cheng, Y., Billin, A.N., Hardison, R.C., Mackay, J.P., et al. (2011). Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. Proceedings of the National Academy of Sciences *108*, E159–E168. 10.1073/pnas.1102140108.
87. Liu, Y., Tavana, O., and Gu, W. (2019). p53 modifications: exquisite decorations of the powerful guardian. J Mol Cell Biol *11*, 564–577. 10.1093/jmcb/mjz060.
88. Louphrasitthiphol, P., Siddaway, R., Loffreda, A., Pogenberg, V., Friedrichsen, H., Schepsky, A., Zeng, Z., Lu, M., Strub, T., Freter, R., et al. (2020). Tuning Transcription Factor Availability through Acetylation-Mediated Genomic Redistribution. Mol Cell *79*, 472-487.e10. 10.1016/j.molcel.2020.05.025.
89. Soutoglou, E., Katrakili, N., and Talianidis, I. (2000). Acetylation Regulates Transcription Factor Activity at Multiple Levels. Mol Cell *5*, 745–751. 10.1016/S1097-2765(00)80253-1.
90. Reed, S., and Quelle, D. (2014). p53 Acetylation: Regulation and Consequences. Cancers (Basel) *7*, 30–69. 10.3390/cancers7010030.
91. Zhang, W., Kadam, S., Emerson, B.M., and Bieker, J.J. (2001). Site-Specific Acetylation by p300 or CREB Binding Protein Regulates Erythroid Krüppel-Like Factor Transcriptional Activity via Its Interaction with the SWI-SNF Complex. Mol Cell Biol *21*, 2413–2422. 10.1128/MCB.21.7.2413-2422.2001.
92. Sun, Z., Chin, Y.E., and Zhang, D.D. (2009). Acetylation of Nrf2 by p300/CBP Augments Promoter-Specific DNA Binding of Nrf2 during the Antioxidant Response. Mol Cell Biol *29*, 2658–2672. 10.1128/MCB.01639-08.
93. Chen, L.-F., and Greene, W.C. (2003). Regulation of distinct biological activities of the NF-?B transcription factor complex by acetylation. J Mol Med *81*, 549–557. 10.1007/s00109-003-0469-0.
94. Tomita, A., Towatari, M., Tsuzuki, S., Hayakawa, F., Kosugi, H., Tamai, K., Miyazaki, T., Kinoshita, T., and Saito, H. (2000). c-Myb acetylation at the carboxyl-terminal conserved domain by transcriptional co-activator p300. Oncogene *19*, 444–451. 10.1038/sj.onc.1203329.
95. Wang, W., Pan, K., Chen, Y., Huang, C., and Zhang, X. (2012). The acetylation of transcription factor HBP1 by p300/CBP enhances p16INK4A expression. Nucleic Acids Res *40*, 981–995. 10.1093/nar/gkr818.
96. Polesskaya, A., Duquet, A., Naguibneva, I., Weise, C., Vervisch, A., Bengal, E., Hucho, F., Robin, P., and Harel-Bellan, A. (2000). CREB-binding Protein/p300 Activates MyoD by Acetylation. Journal of Biological Chemistry *275*, 34359–34364. 10.1074/jbc.M003815200.
97. Zhuang, S. (2013). Regulation of STAT signaling by acetylation. Cell Signal *25*, 1924–1931. 10.1016/j.cellsig.2013.05.007.
98. Wang, C., Tian, L., Popov, V.M., and Pestell, R.G. (2011). Acetylation and nuclear receptor action. J Steroid Biochem Mol Biol *123*, 91–100. 10.1016/j.jsbmb.2010.12.003.
99. Boyes, J., Byfield, P., Nakatani, Y., and Ogryzko, V. (1998). Regulation of activity of the transcription factor GATA-1 by acetylation. Nature *396*, 594–598. 10.1038/25166.

100. Narita, T., Ito, S., Higashijima, Y., Chu, W.K., Neumann, K., Walter, J., Satpathy, S., Liebner, T., Hamilton, W.B., Maskey, E., et al. (2021). Enhancers are activated by p300/CBP activity-dependent PIC assembly, RNAPII recruitment, and pause release. Mol Cell *81*, 2166-2182.e6. 10.1016/j.molcel.2021.03.008.
101. Seto, E., and Yoshida, M. (2014). Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. Cold Spring Harb Perspect Biol *6*, a018713–a018713. 10.1101/cshperspect.a018713.
102. Ververis, K., and Karagiannis, T.C. (2012). An atlas of histone deacetylase expression in breast cancer: fluorescence methodology for comparative semi-quantitative analysis. Am J Transl Res *4*, 24–43.
103. Maniatis, T., Goodbourn, S., and Fischer, J. (1987). Regulation of inducible and tissue-specific gene expression. Science (1979) *236*, 1237–1245. 10.1126/science.3296191.
104. Filtz, T.M., Vogel, W.K., and Leid, M. (2014). Regulation of transcription factor activity by interconnected post-translational modifications. Trends Pharmacol Sci *35*. 10.1016/j.tips.2013.11.005.
105. Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol *16*, 144–154. 10.1038/nrm3949.
106. Li, X.-Y., Thomas, S., Sabo, P.J., Eisen, M.B., Stamatoyannopoulos, J.A., and Biggin, M.D. (2011). The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. Genome Biol *12*, R34. 10.1186/gb-2011-12-4-r34.
107. Iwafuchi-Doi, M., and Zaret, K.S. (2016). Cell fate control by pioneer transcription factors. Development *143*, 1833–1837. 10.1242/dev.133900.
108. Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613–626. 10.1038/nrg3207.
109. Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell *161*, 1012–1025. 10.1016/j.cell.2015.04.004.
110. 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. 10.1016/j.molcel.2013.10.009.
111. Imhof, A., Yang, X.-J., Ogryzko, V. V, Nakatani, Y., Wolffe, A.P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. Current Biology *7*, 689–692. 10.1016/S0960-9822(06)00296-X.
112. Cho, S., Schroeder, S., Kaehlcke, K., Kwon, H.-S., Pedal, A., Herker, E., Schnoelzer, M., and Ott, M. (2009). Acetylation of cyclin T1 regulates the equilibrium between active and inactive P-TEFb in cells. EMBO J *28*, 1407–1417. 10.1038/emboj.2009.99.
113. C. Quaresma, A.J., Bugai, A., and Barboric, M. (2016). Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb. Nucleic Acids Res *44*, 7527–7539. 10.1093/nar/gkw585.

114. Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell *151*, 56–67. 10.1016/j.cell.2012.08.026.
115. Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. Cell *151*, 68–79. 10.1016/j.cell.2012.08.033.
116. Faiola, F., Liu, X., Lo, S., Pan, S., Zhang, K., Lymar, E., Farina, A., and Martinez, E. (2005). Dual Regulation of c-Myc by p300 via Acetylation-Dependent Control of Myc Protein Turnover and Coactivation of Myc-Induced Transcription. Mol Cell Biol *25*, 10220–10234. 10.1128/MCB.25.23.10220-10234.2005.
117. Vervoorts, J., Lüscher-Firzlaff, J.M., Rottmann, S., Lilischkis, R., Walsemann, G., Dohmann, K., Austen, M., and Lüscher, B. (2003). Stimulation of c-MYC transcriptional activity and acetylation by recruitment of the cofactor CBP. EMBO Rep *4*, 484–490. 10.1038/sj.embor.embor821.
118. Anderson, E., Devenney, P.S., Hill, R.E., and Lettice, L.A. (2014). Mapping the Shh long-range regulatory domain. Development *141*, 3934–3943. 10.1242/dev.108480.
119. Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., Mach, P., Kryzhanovska, M., Tihanyi, G., Kohler, H., et al. (2021). Nonlinear control of transcription through enhancer-promoter interactions. . bioRxiv. https://doi.org/10.1101/2021.04.22.440891.
120. Kane, L., Williamson, I., Flyamer, I.M., Kumar, Y., Hill, R.E., Lettice, L.A., and Bickmore, W.A. (2021). Cohesin is required for long-range enhancer action. . bioRxiv.
121. McSwiggen, D.T., Hansen, A.S., Teves, S.S., Marie-Nelly, H., Hao, Y., Heckert, A.B., Umemoto, K.K., Dugast-Darzacq, C., Tjian, R., and Darzacq, X. (2019). Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. Elife *8*. 10.7554/eLife.47098.
122. Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer–promoter contacts in gene expression control. Nat Rev Genet *20*, 437–455. 10.1038/s41576-019-0128-0.
123. Bock, A., Annibale, P., Konrad, C., Hannawacker, A., Anton, S.E., Maiellaro, I., Zabel, U., Sivaramakrishnan, S., Falcke, M., and Lohse, M.J. (2021). Optical Mapping of cAMP Signaling at the Nanometer Scale. Cell *184*, 2793. 10.1016/j.cell.2021.04.043.
124. Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. Curr Opin Genet Dev *22*, 148–155. 10.1016/j.gde.2012.02.013.
125. Näär, A.M., Lemon, B.D., and Tjian, R. (2001). Transcriptional Coactivator Complexes. Annu Rev Biochem *70*, 475–501. 10.1146/annurev.biochem.70.1.475.
126. Bulynko, Y.A., and O'Malley, B.W. (2011). Nuclear Receptor Coactivators: Structural and Functional Biochemistry. Biochemistry *50*, 313–328. 10.1021/bi101762x.
127. Lewis, B.A., and Reinberg, D. (2003). The mediator coactivator complex: functional and physical roles in transcriptional regulation. J Cell Sci *116*, 3667–3675. 10.1242/jcs.00734.
128. Schmidt, S.F., Larsen, B.D., Loft, A., and Mandrup, S. (2016). Cofactor squelching: Artifact or fact? BioEssays *38*, 618–626. 10.1002/bies.201600034.
129. Vo, N., and Goodman, R.H. (2001). CREB-binding Protein and p300 in Transcriptional Regulation. Journal of Biological Chemistry *276*, 13505–13508. 10.1074/jbc.R000025200.

130. Hatazawa, S., Liu, J., Takizawa, Y., Zandian, M., Negishi, L., Kutateladze, T.G., and Kurumizaka, H. (2022). Structural basis for binding diversity of acetyltransferase p300 to the nucleosome. iScience *25*, 104563. 10.1016/j.isci.2022.104563.
131. Kikuchi, M., Morita, S., Wakamori, M., Sato, S., Uchikubo-Kamo, T., Suzuki, T., Dohmae, N., Shirouzu, M., and Umehara, T. (2023). Epigenetic mechanisms to propagate histone acetylation by p300/CBP. Nat Commun *14*, 4103. 10.1038/s41467-023-39735-4.
132. Zhang, Y., Brown, K., Yu, Y., Ibrahim, Z., Zandian, M., Xuan, H., Ingersoll, S., Lee, T., Ebmeier, C.C., Liu, J., et al. (2021). Nuclear condensates of p300 formed though the structured catalytic core can act as a storage pool of p300 with reduced HAT activity. Nat Commun *12*, 4618. 10.1038/s41467-021-24950-8.
133. Zhang, Y., Xue, Y., Shi, J., Ahn, J., Mi, W., Ali, M., Wang, X., Klein, B.J., Wen, H., Li, W., et al. (2018). The ZZ domain of p300 mediates specificity of the adjacent HAT domain for histone H3. Nat Struct Mol Biol *25*, 841–849. 10.1038/s41594-018-0114-9.
134. Ma, L., Gao, Z., Wu, J., Zhong, B., Xie, Y., Huang, W., and Lin, Y. (2021). Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. Mol Cell *81*, 1682-1697.e7. 10.1016/j.molcel.2021.01.031.
135. Zeng, L., Zhang, Q., Gerona-Navarro, G., Moshkina, N., and Zhou, M.-M. (2008). Structural Basis of Site-Specific Histone Recognition by the Bromodomains of Human Coactivators PCAF and CBP/p300. Structure *16*, 643–652. 10.1016/j.str.2008.01.010.
136. Ma, L., Gao, Z., Wu, J., Zhong, B., Xie, Y., Huang, W., and Lin, Y. (2021). Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. Mol Cell *81*, 1682-1697.e7. 10.1016/j.molcel.2021.01.031.
137. Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., et al. (2012). Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. Cell *149*, 214–231. 10.1016/j.cell.2012.02.013.
138. Heckert, A., Dahal, L., Tjian, R., and Darzacq, X. (2022). Recovering mixtures of fast-diffusing states from short single-particle trajectories. Elife *11*. 10.7554/eLife.70169.
139. Chen, Y., Cattoglio, C., Dailey, G.M., Zhu, Q., Tjian, R., and Darzacq, X. (2022). Mechanisms governing target search and binding dynamics of hypoxia-inducible factors. Elife *11*. 10.7554/eLife.75064.
140. de Jonge, W.J., Patel, H.P., Meeussen, J.V.W., and Lenstra, T.L. (2022). Following the tracks: How transcription factor binding dynamics control transcription. Biophys J *121*, 1583–1592. 10.1016/j.bpj.2022.03.026.
141. Liu, X., Wang, L., Zhao, K., Thompson, P.R., Hwang, Y., Marmorstein, R., and Cole, P.A. (2008). The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. Nature *451*, 846–850. 10.1038/nature06546.
142. Lasko, L.M., Jakob, C.G., Edalji, R.P., Qiu, W., Montgomery, D., Digiammarino, E.L., Hansen, T.M., Risi, R.M., Frey, R., Manaves, V., et al. (2017). Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature *550*, 128–132. 10.1038/nature24028.

143. Black, J.C., Choi, J.E., Lombardo, S.R., and Carey, M. (2006). A Mechanism for Coordinating Chromatin Modification and Preinitiation Complex Assembly. Mol Cell *23*, 809–818. 10.1016/j.molcel.2006.07.018.
144. Wang, F., Marshall, C.B., and Ikura, M. (2013). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. Cellular and Molecular Life Sciences *70*, 3989–4008. 10.1007/s00018-012-1254-4.
145. Berlow, R.B., Dyson, H.J., and Wright, P.E. (2017). Hypersensitive termination of the hypoxic response by a disordered protein switch. Nature *543*, 447–451. 10.1038/nature21705.
146. Teufel, D.P., Bycroft, M., and Fersht, A.R. (2009). Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2. Oncogene *28*, 2112–2118. 10.1038/onc.2009.71.
147. Wunderlich, Z., and Mirny, L.A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. Trends in Genetics *25*, 434–440. 10.1016/j.tig.2009.08.003.
148. Teufel, D.P., Freund, S.M., Bycroft, M., and Fersht, A.R. (2007). Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. Proceedings of the National Academy of Sciences *104*, 7009–7014. 10.1073/pnas.0702010104.