Trends in Biochemical Sciences

CellPress
REVIEWS

Opinion

# 'Nonlinear' Biochemistry of Nucleosome Detergents

Alexandre M. Erkine[1,*]

**The transcriptional activation domains (TADs) are critical for life, yet intrinsically disordered polypeptides with no specific consensus sequence, interacting with multiple targets via low-specificity fuzzy contacts. The recent integration of machine learning approaches in biochemistry allows analysis of large experimental datasets of functional TADs as a whole and clear observation of TAD features. The emerging picture describes TADs as sequences without consensus but with a variety of detergent-like mini-motifs enriched in negatively charged and aromatic amino acids. Comparison of the canonical direct coactivator recruitment model and a new model describing TADs as nucleosome detergents that trigger chromatin remodeling during gene activation helps solve a fundamental enigma of molecular biology spanning 30 years.**

## Highlights

Introduction of machine learning in analyzing TADs led to the realization that the main feature of TAD sequences is presence of a variety of detergent-like mini-motifs comprising negatively charged and aromatic extremities separated by variable spacers.

Interchangeability of TADs, absence of a discernable consensus sequence, and high frequency (1/100) of functional TADs appearance in random sequence pools suggests that TADs operate at the near-stochastic level of interactions.

While approximation of interactions to the near-stochastic level renders the direct coactivator recruitment model for TADs unworkable, it fits perfectly with the gene promoter nucleosome detergent model.

Acceptance of functionality at near-stochastic levels is challenging for canonical biochemistry and requires a drastic revision of mentality and methodology.

## Paradigm Shift in Traditional Biochemistry

A classical biochemical understanding of cellular processes is based on specific molecular structures interacting with a high level of specificity and affinity to fulfill vital cellular functions. However, this paradigm has recently undergone a shift with the realization that a large if not dominant fraction of proteins and protein regions in the crowded tiny cellular space, are intrinsically disordered, yet perfectly functional [1], and that interactions between biological molecules are fuzzy [2], often with low specificity and affinity to the extent of being near-stochastic. This shift presents a formidable challenge to classical biochemistry as interactions at the near-noise level are traditionally considered as functionally nonessential, with the majority of current biochemical methods obligatorily discarding **near-stochastic interactions** (see Glossary).

**Transcriptional activation domains** (TADs) are a typical example of **intrinsically disordered regions (IDRs)** that perform a vital function with a mechanism that remains vague despite numerous attempts for clarification over more than 30 years [3–7]. This extended period of ambiguity indicates the deficiency of traditional biochemical approaches, and the need for a fundamental shift in mentality and methodology to understand the mechanisms of TAD functionality.

## Classical View on Mechanisms of TADs Functionality

Since their initial discovery more than 30 years ago, direct physical recruitment of basal transcriptional machinery components and coactivators to gene promoters was the proposed mechanism for TAD function [8] (Figure 1). The number of potential interacting targets of TADs grew progressively with the discovery of new players in transcription regulation. Over the years these interactions were demonstrated with basal transcriptional machinery components such as TBP [9–13], TFIIB [14,15], TFIIH [16,17], TFIIA [18,19], RNA polymerase II [20], variable TAFs [21], Mediator subunits Med17 (Srb4), Srb10, Med15 (Gal11), Med2, and Med 25 [5,7,22–27], as well as with subunits of chromatin remodeling and histone-modifying complexes such as Ada2, Taf17, Tra1 (SAGA and NuA4 complexes) [28–34], Swi1 and Snf2 (SWI/SNF complex)

[1]Butler University, Indianapolis, IN 46208, USA

*Correspondence: aerkine@butler.edu (A.M. Erkine).

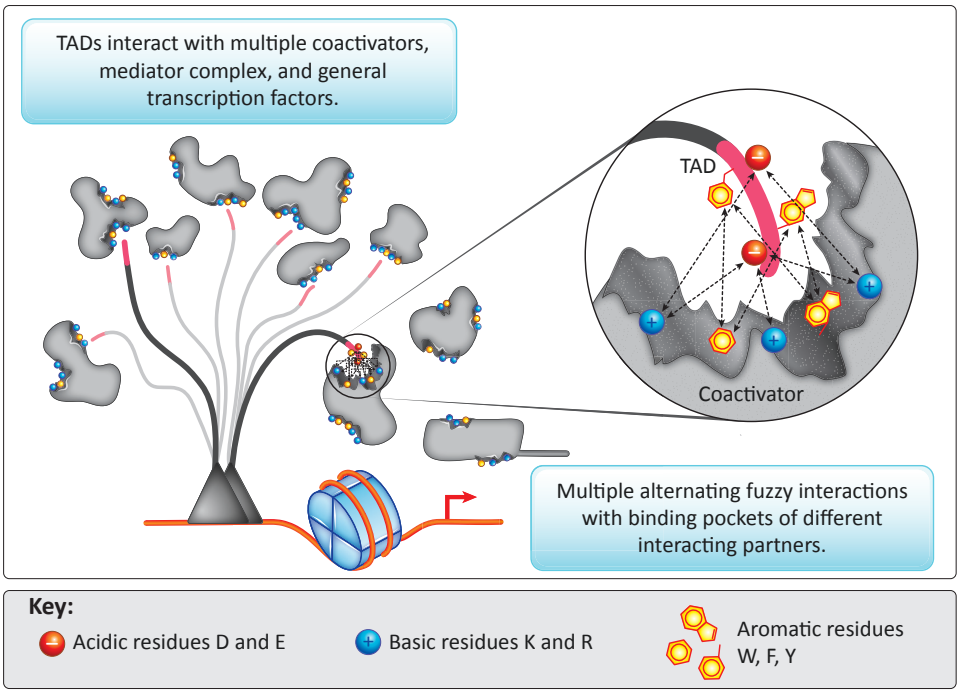


Figure 1. Direct Recruitment Model. Gene activators are represented by gray triangles (DNA-binding domains) with black linkers connected to pink TADs at the end. Two (or more) TADs are simultaneously involved in different low specificity fuzzy interactions with multiple coactivators, Mediator, and basal transcription factors (gray shapes) necessary for chromatin remodeling and transcription initiation complex assembly. Interactions happen in parallel and in line with the liquid–liquid phase separation model [54]. TADs are enriched in acidic and aromatic residues that interact with basic and aromatic pockets within the interaction targets (circular inset of magnifying glass). Nucleosomes with blue cylinder representing histone octamer provide a chromatin context unique for each promoter. Elements of the picture are adopted from publications [56] and [7]. Abbreviations: TAD, transcriptional activation domain.

[34–36], and CBP [37–39]. The large number of potential TAD-interacting targets was estimated to be >200 [40] with no clear ranking system.

Investigation of TAD sequences revealed that they are interchangeable with retention of functionality between activators within a specific eukaryotic cell, and between cells belonging to different biological phyla such as plants, fungi, and mammals [8,41–45]. Despite this functional conservation, there is little conservation in their amino acid sequence, structure, length (for review see [3,46]), and even chemical nature, as native TADs can be replaced with functionally active sequences of non-natural d-amino acids [47], stretches of RNA [48], or small isoxazolidines type compounds [49,50]. The natural TAD sequences can be roughly grouped into acidic, glutamine-rich, and proline-rich categories (for review see [3,46]). However, in most cases, including non-natural compounds [49,50], TADs have negatively charged and hydrophobic extremities.

The multiplicity and variability of potential TAD-interacting targets combined with the uncertainties of TAD composition and structure creates a number of complex mechanistic conundrums. For instance: (i) it is not clear if any given activator molecule has a specific preferred target, and whether there are preference criteria in case of target selection; (ii) it is not clear how, despite the low specificity of interactions, TADs efficiently select and physically

## Glossary

**Activator binding domains (ABDs):** protein regions that interact with TADs. ABDs are present in coactivators including chromatin-remodeling and histone-modifying complexes, general transcription factors, and subunits of the transcriptional Mediator complex.
**Direct recruitment:** a predominant text book mechanism explaining the function of transcriptional activators. Through direct sustained physical contacts with TADs, coactivators and other transcription related components are brought by activators to gene promoters for transcription initiation.
**Intrinsically disordered regions (IDRs):** malleable parts of proteins that do not adopt any specific structure. Almost half of all proteins are either fully disordered or contain IDRs.
**Near-stochastic interactions:** interactions between proteins characterized by extremely high Kd ($10^{-1}$–$10^{-2}$). These interactions are traditionally considered as random, non-functional, and thus unimportant. In a majority of canonical biochemical methods, these interactions are obligatorily discarded as detrimental noise.
**Nonlinear biochemistry:** the term nonlinear is borrowed from physics and mathematics, where it describes multifactorial, highly chaotic processes of real life such as turbulence of high velocity streams, climatic weather pattern developments, movement of objects with a flexible shape through air and water, etc.
**Transcriptional activation domains (TADs):** present in the majority of eukaryotic gene transcription factors also known as activators. TADs determine whether transcription is initiated and at what level.

recruit specific enzymatic complexes necessary for promoter nucleosome evictions, often within seconds after the initial stimulus [51,52]; and (iii) it is not clear how the same TAD sequence can retain a comparable level of functionality when transferred between evolutionarily distant contexts of mammalian, plant, and fungal cells [8,41–45]. The list of ambiguities can be extended further if more detailed examples of TADs are considered, but a key concern is the questionable compatibility of the low or very low TAD sequence-structure specificity, with the ability to choose and recruit a relevant TAD target to a specific gene promoter.

The uncertainties with potential TAD-interacting targets and definitions of TAD sequence features led to postulation of the fuzzy interactions concept in transcription regulation and beyond [1,2,53]. According to this concept, during transcription activation a specific TAD is engaged in multiple low-affinity and low-specificity interactions with a potential target. For instance, in the best-documented case, two TAD regions of Gcn4 are intermittently engaged in fuzzy interactions with three **activator binding domains (ABDs)** and an additional KIX domain of Med15–a subunit of the large Mediator complex [7]. Each ABD in this case forms a center with an increased concentration of positively charged and hydrophobic extremities, allowing longer retention of the negatively charged and hydrophobic TADs in this pocket (Figure 1), where it alternates between low-affinity sites, thus increasing the probability of bond formation required for physical recruitment of the entire Mediator complex. The functionality of these interactions in the case of Gcn4 is reinforced by the correlation of *in vitro* effects generated by specific Gcn4 mutations affecting interactions with Med15 and the *in vivo* activities of Gcn4 carrying the same mutations [5]. Similar correlations were demonstrated between the functionality of VP16 TAD interactions with Tra1, a large subunit of SAGA and Nua4 complexes [28–30].

Fuzzy interactions of the TAD–ABD capturing mechanism partially alleviate the specificity problem by a several-fold increase in the effective concentration of interacting points in the ABD pockets, although the multiplicity of potential TAD targets in the nucleus still creates a largely chaotic environment (Figure 1). The recently formulated liquid–liquid phase separation (LLPS) model for transcriptional control might introduce relative order into this picture [54]. According to this model, gene activators and coactivators coalesce in specific phase-separated nuclear compartments, thus creating transcriptional superenhancers. Using the super-enhancer model to solve the TAD-target specificity problem is intriguing, but was not attempted in the original study, possibly because concentration of greatly variable TADs in a specific center and the multiplicity of possible interacting targets with variable ABDs contributes to the aforementioned recruitment specificity problem. Moreover, the regulatory mechanisms for individual genes entry and exit to and from these superenhancers are not clear.

Summarizing the current state of our understanding of TAD functionality, it is commonly accepted that the mechanism involves fuzzy low-specificity interactions with a large number of possible targets. Organizing the multiplicity of these recruitment events into an orderly fashion remains largely challenging. The newly developed high-throughput approaches [55,56] including bioinformatics and machine learning (ML) open a different perspective, substituting the analysis of ‘trees’ of individual TAD mutations with analysis of the ‘forest’ of all natural and synthetic TADs.

## TADs as Near-Stochastic Nucleosome Detergents

The recent analysis of 6340 mutant variants of the 44-amino-acid-long cAD of Gcn4 [55] revolved around three known TAD features: negative charge, hydrophobicity, and intrinsic disorder. Mutation sets systematically altering each of these features individually revealed that within the scope of performed experiments, the combination of all three features contributes to

TAD functionality, suggesting that the TAD sequence is maintained in an unfolded state by a number of repelling negatively charged amino acids interspersed with hydrophobic, primarily aromatic, residues. The central WxxLF motif of this TAD plays a critical role in interactions with the Med15 subunit of the Mediator complex. Although these mechanistic conclusions are not novel [5,7], this study is a proof of principle for the high-throughput approach analyzing large mutant sets, simultaneously testing different hypotheses or TAD features. A large set of mutants allows an ML approach based on a support vector regression (SVR) ML model for predicting functionality [55], although training the model in this case was restricted by the near absence in the pool of nonfunctional sequences and by a general bias (68%) toward sequences more active than the wild type.

A more comprehensive ML approach was used when analyzing an unbiased 67 263 units pool of random sequences individually replacing the C-terminal TAD of HSF1 activator [56]. Utilization of several ML models such as Lasso, Ridge, and Xgboost allowed comparison of these models in prediction of TAD functionality and showed good correlation. The ML model training was performed using both functional and nonfunctional subpools. Additionally, these non-black-box models, while using a large number of different sequence features (>100), allow calculation of their importance for each individual feature in model performance. The outcomes of this ML analysis reveal that the most important features describing functional TADs are the enrichment of aromatic and negatively charged amino acids, with a single feature that described redundant mini-motifs devoid of lysine and arginine, and contained only aromatic and negatively charged amino acids, showing 74% of the ML gain in the Lasso model [56]. Similar outcomes emphasizing a mere presence of negative charge and aromaticity appeared after analyzing yet another TAD library containing 962 mutant variants of 13 known natural TADs.

Both recent investigations of large TAD sequence pools [55,56] are consistent with the recently formulated model for TAD functionality as detergents [46] – sequences with hydrophobic (aromatic) and hydrophilic (negatively charged) extremities separated by a varying length spacer. The outcomes of random sequence library screening [56] describing aromatic and negatively charged mini-motifs are consistent with this description. The nucleosome detergent model [46] postulates that negatively charged and aromatic mini-motifs interact with abundantly represented nucleosomes using the negatively charged extremity to interact with the lysines and arginines on histones, while using the aromatic extremity to intercalate between DNA bases, thus peeling off the nucleosomal DNA, or alternatively, interacting with positively charged hydrophobic pockets of the histone core to distort the nucleosome structure, thus triggering gene promoter chromatin remodeling (Figure 2). Involvement of a limited number of contacts suggests low affinity and specificity in these interactions and a large number of possible fuzzy interaction sites. Although interactions with any nucleosome to multiple sites at stochastic or near-stochastic level are possible, interactions with specific promoter nucleosomes are increased by many orders of magnitude due to anchoring of TAD by the DNA-binding domain of the activator, thus significantly increasing the effective local concentration of TADs (Figure 2 and Box 1).

Both the **direct recruitment** and nucleosome detergent models operate with fuzzy low-affinity low-specificity interactions, however how low is the specificity, considered in these two models very differently. Going to the extreme of near-stochastic interactions makes the direct recruitment model decreasingly workable (bottom level of Kd is $\sim 10^{-5}$ [7]), whereas the nucleosome detergent model actually requires this near-randomness (Kd $\sim 10^{-1}$–$10^{-2}$) for functionality, as more specific interactions lead either to the sequestration of this TAD in a specific location, or to promoter nucleosome immobilization by the TAD, thus resulting in transcription repression.

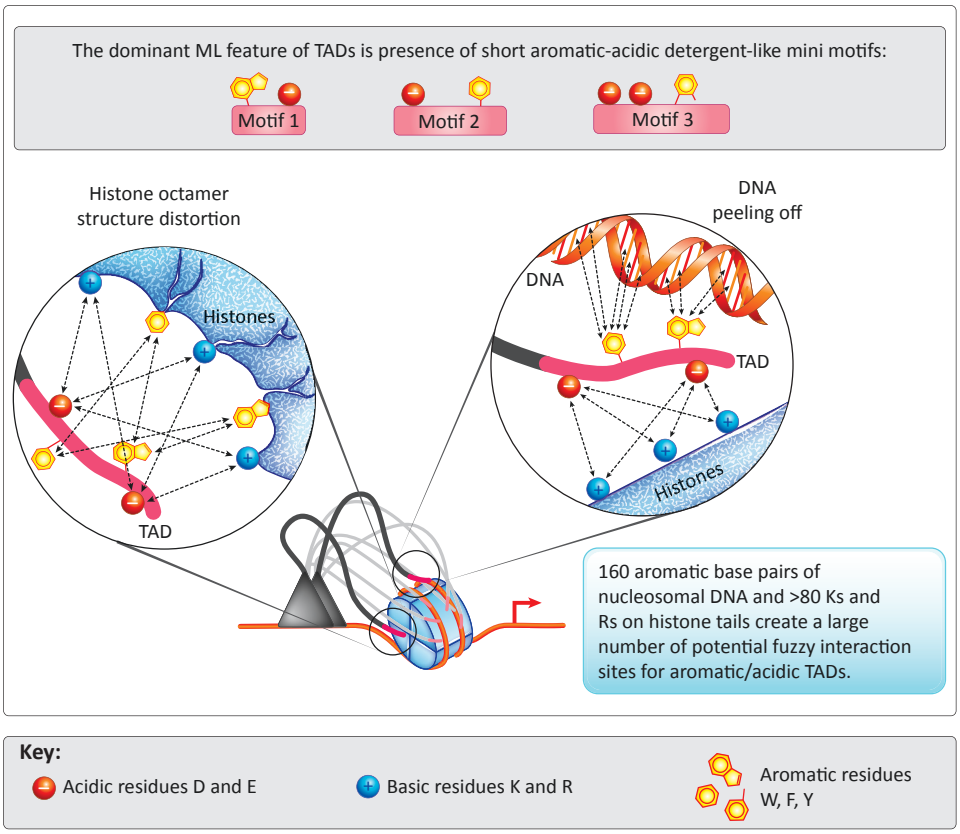


Figure 2. Nucleosome Detergent Model. Designation of shapes is similar to Figure 1. The promoter nucleosome has multiple sites enriched in basic and aromatic residues. The main source of aromaticity is the DNA bases. TADs are involved in multiple near-stochastic fuzzy interactions forming transient electrostatic and aromatic ($\pi$–$\pi$) bonds, thus peeling off the DNA (circular inset of magnifying glass to the right) or simply distorting the histone octamer structure (circular inset of magnifying glass to the left) and triggering chromatin remodeling as a step preceding the recruitment of transcription machinery (Figure 1). Post-translational modifications of TADs, for example, phosphorylation or acetylation that brings additional negative charge, might regulate broader chromatin remodeling and formation of liquid–liquid phase separation [54] bringing multiple promoters together into a transcriptosome. Abbreviations: ML, machine learning; TAD, transcriptional activation domain.

The outcome that ~1% (739 sequences) of the random sequences pool (67,263) represents functional TADs [56] suggests that TADs operate at the extremely low near-stochastic level of specificity. Similar outcomes of 0.1–2% of random sequence pools being functional were demonstrated for screens performed in contexts of other activators such as Gal4 [57] and *lexA* hybrids [58]. Considering the size of the random HSF library of TADs with ≤20 amino acid

## Box 1. High and Low Specificity in One Molecule

An important point for the functionality of regions interacting with targets at the near-stochastic level is their intramolecular fusion to another domain that is characterized by high-specificity interactions. In case of a TAD, it is fused to the DNA-binding domain (DBD) of the activator. The DBD anchors its TAD to the cognate gene promoter, thus increasing the local concentration of TAD by many orders of magnitude (inversely proportional to the DBD Kd value), and catalyzing interactions of the TAD and local nucleosomes. Without high affinity anchoring, TADs interact stochastically with the nucleosomal continuum of the nucleus.

positions [56], which is an approximation for the functional TAD length [59], with all 20 natural amino acids for the sequence, the theoretical combinatorial space is $20^{20}$ with $1\% \approx 10^{24}$ of potential sequence combinations, all able to work within the context of the same activator molecule. The number $10^{24} \approx 16^{20}$, which is the number of all possible amino acid combinations for a stretch of 20 positions if only four letters (e.g. K, R, and two others) are removed from the amino acid alphabet. These considerations, the astronomical number of variations, and short size of TADs with few critical residues, imply that almost any protein might have a TAD-like module, suggesting a close approximation to the stochastic level of interactions. Consistent with the stochastic nature of interactions is the fact that almost all 6340 sequence variations for the near-saturation mutagenesis of the GCN4 cTAD are transcriptionally active [55]. The nucleosome detergent model easily operates at the near-stochastic level and explains functionality at the wild-type level of seemingly monotonous sequences such as WDWDWDWDWDWD [56] or FDFDFDFDFDFD (same study unpublished data), which contain a combination of redundant detergent-like mini-motifs, whereas the direct recruitment model requires a consensus sequence such as the 9-amino-acid consensus [44,60] or WxxLF [7]. Neither of these motifs was found to dominate the random sequence subpool of functional TADs and both were over-represented among the nonfunctional sequences [56]. Additionally, the 9-amino-acid consensus sequence failed when used as a feature for ML functionality prediction in the random sequence library [56], while performing decently as a feature in the narrow library of natural TADs from which this loose consensus was derived [44,60].

The direct recruitment and nucleosome detergent models are not necessarily mutually exclusive and likely coexist. The low affinity of TAD recruitment interactions with transcription preinitiation complex (PIC) including Mediator is likely boosted by cooperative interactions of the PIC complex with nucleosome-free core promoter elements such as TATA box and others. This recruitment and stabilization of PIC components by TADs was demonstrated in early *in vitro* experiments on nucleosome-free promoters (reviewed in [8]). Both the direct recruitment and nucleosome detergent model are compatible with the LLPS model [54] (Figures 1 and 2 legends). The nucleosome detergent model has an advantage of being able to explain mechanisms of the phase separation initiation. The well-known inducible and reversible post-translational modifications of TADs, such as phosphorylation and acetylation, by bringing negative charges, create additional acidic–aromatic detergent-like amino acid combinations leading to more efficient chromatin remodeling. Large scale chromatin remodeling can easily promote coalescence into a super enhancer or a transcription center containing multiple gene promoters forming a phase separated transcriptosome.

Acceptance of the functional importance of near-stochastic interactions is challenging for canonical biochemistry because these interactions are traditionally considered mechanistically nonessential and are obligatorily discarded by standard biochemical methods such as pull-downs, crosslinking, Y2H, FRET, ChIP, etc. Newly developed modifications of methods including NMR for fuzzy interactions [61,62] reflect the occurring shift in classical biochemistry mentality toward acceptance of weak dynamic interactions. The inapplicability of standard methods in the case of TADs likely explains the unusually long (>30 years) struggle to understand the mechanism of TAD functionality [3–7]. Importantly, the extreme of near-stochastic fuzzy interactions that makes the nucleosome detergent model theoretically workable, suggests inapplicability of canonical biochemistry in attempts to prove or disprove it, and requires an initial fundamental overturn of mentality and development of new methodological tools (Box 2).

The development of ML approaches based on modern statistics allows better detection of near-stochastic interactions in the crowded nuclear space and clarification of the features of a

### Box 2. Drastic Change in Mentality Is Obligatory

A typical demand from an established biochemist is: prove the model by designing an *in vitro* experiment demonstrating nucleosome distortion/relocation by a TAD. The problem with this approach is that the near-stochastic contacts of a TAD (nucleosome detergent) at multiple points of the nucleosome (e.g., intercalation of an aromatic TAD residue between DNA bases and interactions with positively charged Ks and Rs of histones) are highly transient with extremely minute effects for detection using conventional biochemistry. Additionally, these multiple transient distortions of the nucleosome likely only trigger promoter chromatin remodeling, which is actually mediated by enzymatic activities of chromatin remodeling complexes. Adding these complexes to the *in vitro* experiment mix flips the mentality towards the familiar direct recruitment model.

### Box 3. "Nonlinear" biochemistry

While the trajectory, travel time, and landing point are easily and accurately calculated for a metal ball, it is impossible to do so for a paper plane or a tree leaf. This nonlinearity of the real world (think weather forecasts) is dealt (often with help of ML) by nonlinear branches of older sciences such as physics and mathematics. Currently, the field of intrinsically disordered regions functioning at the near-stochastic level of interactions in an extremely crowded cellular space is likely to morph into a **nonlinearbranch of biochemistry**.

TAD 'forest' [56] instead of concentrating on individual mutational 'trees' of specific TADs. This maturation of biochemistry is akin to the development of nonlinear concepts for noisy multifactorial processes in older sciences such as physics and mathematics (Box 3). Acceptance of functionality at the near-stochastic level of interactions between biological molecules connected to the randomness of the noisy real-world environment while solving old enigmas opens many doors for new discoveries.

## Concluding Remarks

Currently, acceptance of functionality at the near-stochastic level of interactions (Kd in the $10^{-1}$–$10^{-2}$ interval) is the most challenging concept. This nonlinear biochemical perspective deems canonical biochemistry – both mentality and methodology – inapplicable and demands flexibility toward new hypotheses and revision of mechanistic models. TADs remain enigmatic because they operate at the near-stochastic level. Utilization of machine learning based on analysis of wet-lab-generated datasets is a promising development allowing observation of feature modifications for the whole TADs 'forest' in response to changes in the intramolecular activator or gene-reporter contexts [55,56], and to intracellular changes associated with genome modifications (see Outstanding Questions).

### Acknowledgments

I am grateful to Charles N.J. Ravarani for numerous discussions, ML insights, and critical suggestions, and to Daniel C. Dudman, and Bradley K. Broyles for help with manuscript preparation. This study was supported by the Butler University Innovation Grant and a grant from the National Science Foundation (MCB-1029254).

### References

1. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* 83, 553–584
2. Fuxreiter, M. and Tompa, P. (2012) Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.* 725, 1–14
3. Mapp, A.K. and Ansari, A.Z. (2007) A TAD further: exogenous control of gene activation. *ACS Chem. Biol.* 2, 62–75
4. Hahn, S. and Young, E.T. (2011) Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* 189, 705–736
5. Warfield, L. *et al.* (2014) A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci. U. S. A.* 111, E3506–3513
6. Bhaumik, S.R. and Green, M.R. (2001) SAGA is an essential *in vivo* target of the yeast acidic activator Gal4p. *Genes Dev.* 15, 1935–1945
7. Tuttle, L.M. *et al.* (2018) Gcn4-Mediator specificity is mediated by a large and dynamic fuzzy protein–protein complex. *Cell Rep.* 22, 3251–3264
8. Ptashne, M. and Gann, A. (1997) Transcriptional activation by recruitment. *Nature* 386, 569–577

## Outstanding Questions

Switching from canonical biochemistry of specific or semispecific interactions to the multiplicity of fuzzy near-stochastic interactions reality, the ML approaches become invaluable. Considering a possibility of creating large libraries of synthetic TAD sequences for an individual activator/reporter, screening them *in vivo* (microevolution), and analyzing large functional subpools as a whole using ML, a number of interesting questions become apparent.

What TAD features become critical if: (i) the reporter gene promoter is nucleosome free or has strongly positioned nucleosomes? (ii) Core promoter sequence elements are modified? (iii) The activator molecule is engineered to be short-lived due to degradation (fast acting TADs)? (iv) The cellular environment is modified by genome modifications (e.g., knockout of genes coding for chromatin remodeling and histone modifying activities, histone chaperones, Mediator complex components, and general transcription factors)?

While modifying intramolecular and cellular contexts for TADs, does it potentially shift the requirement for higher or lower specificity, thus shifting the equilibrium toward more specific sequences expected by the direct recruitment model or toward sequence pools with less defined and variable TADs?

In all analyzed TAD pools there is a small fraction of functional sequences that represents features opposite to the general rule (e.g., sequences with net positive charge). What makes these exceptions functional?

What are specific TAD features that render them nonfunctional (e.g., presence of specific sequence motifs)?

By developing precise ML models and utilizing CRISP-Cas9 technologies, can we predict and correspondingly edit natural TADs, thus modulating the functionality level?

9. Horikoshi, M. *et al.* (1988) Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54, 1033–1042
10. Stringer, K.F. *et al.* (1990) Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* 345, 783–786
11. Hermann, S. *et al.* (2001) How transcriptional activators bind target proteins. *J. Biol. Chem.* 276, 40127–40132
12. Capella, M. *et al.* (2014) Plant homeodomain-leucine zipper I transcription factors exhibit different functional AHA motifs that selectively interact with TBP or/and TFIIB. *Plant Cell Rep.* 33, 955–967
13. Khan, S.H. *et al.* (2011) TBP binding-induced folding of the glucocorticoid receptor AF1 domain facilitates its interaction with steroid receptor coactivator-1. *PLoS One* 6, e21939
14. Lin, Y.S. *et al.* (1991) Binding of general transcription factor TFIIB to an acidic activating region. *Nature* 353, 569–571
15. Choy, B. and Green, M.R. (1993) Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* 366, 531–536
16. Xiao, H. *et al.* (1994) Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. *Mol. Cell. Biol.* 14, 7013–7024
17. Chabot, P.R. *et al.* (2014) Structural and functional characterization of a complex between the acidic transactivation domain of EBNA2 and the Tfb1/p62 subunit of TFIIH. *PLoS Pathog.* 10, e1004042
18. Stargell, L.A. and Struhl, K. (1995) The TBP-TFIIA interaction in the response to acidic activators *in vivo*. *Science* 269, 75–78
19. Ozer, J. *et al.* (1996) Transcription factor IIA mutations show activator-specific defects and reveal a IIA function distinct from stimulation of TBP-DNA binding. *J. Biol. Chem.* 271, 11182–11190
20. Tan, Q. *et al.* (2000) Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. *Genes Dev.* 14, 339–348
21. Goodrich, J.A. and Tjian, R. (1994) TBP–TAF complexes: selectivity factors for eukaryotic transcription. *Curr. Opin. Cell Biol.* 6, 403–409
22. Koh, S.S. *et al.* (1998) An activator target in the RNA polymerase II holoenzyme. *Mol. Cell* 1, 895–904
23. Myers, L.C. *et al.* (1999) Mediator protein mutations that selectively abolish activated transcription. *Proc. Natl. Acad. Sci. U. S. A.* 96, 67–72
24. Jeong, C.J. *et al.* (2001) Evidence that Gal11 protein is a target of the Gal4 activation domain in the mediator. *Biochemistry* 40, 9421–9427
25. Ansari, A.Z. *et al.* (2002) Transcriptional activating regions target a cyclin-dependent kinase. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14706–14709
26. Qiu, H. *et al.* (2004) An array of coactivators is required for optimal recruitment of TATA binding protein and RNA polymerase II by promoter-bound Gcn4p. *Mol. Cell. Biol.* 24, 4104–4117
27. Aguilar, X. *et al.* (2014) Interaction studies of the human and *Arabidopsis thaliana* Med25-ACID proteins with the herpes simplex virus VP16- and plant-specific Dreb2a transcription factors. *PLoS One* 9, e98575
28. Knutson, B.A. and Hahn, S. (2011) Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Mol. Cell. Biol.* 31, 818–831
29. Brown, C.E. *et al.* (2001) Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* 292, 2333–2337
30. Bhaumik, S.R. *et al.* (2004) *In vivo* target of a transcriptional activator revealed by fluorescence resonance energy transfer. *Genes Dev.* 18, 333–343
31. Barlev, N.A. *et al.* (1995) Characterization of physical interactions of the putative transcriptional adaptor, ADA2, with acidic activation domains and TATA-binding protein. *J. Biol. Chem.* 270, 19337–19344
32. Thut, C.J. *et al.* (1995) p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. *Science* 267, 100–104
33. Henriksson, A. *et al.* (1997) Role of the Ada adaptor complex in gene activation by the glucocorticoid receptor. *Mol. Cell. Biol.* 17, 3065–3073
34. Neely, K.E. *et al.* (1999) Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. *Mol. Cell* 4, 649–655
35. Neely, K.E. *et al.* (2002) Transcription activator interactions with multiple SWI/SNF subunits. *Mol. Cell. Biol.* 22, 1615–1625
36. Prochasson, P. *et al.* (2003) Targeting activity is required for SWI/SNF function *in vivo* and is accomplished through two partially redundant activator-interaction domains. *Mol. Cell* 12, 983–990
37. Kundu, T.K. *et al.* (2000) Activator-dependent transcription from chromatin *in vitro* involving targeted histone acetylation by p300. *Mol. Cell* 6, 551–561
38. Lau, O.D. *et al.* (2000) HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol. Cell* 5, 589–595
39. Mukherjee, S.P. *et al.* (2013) Analysis of the RelA:CBP/p300 interaction reveals its involvement in NF-kappaB-driven transcription. *PLoS Biol.* 11, e1001647
40. Keung, A.J. *et al.* (2014) Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell* 158, 110–120
41. Beerli, R.R. *et al.* (1998) Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14628–14633
42. Escher, D. *et al.* (2000) Conservation of glutamine-rich transactivation function between yeast and humans. *Mol. Cell. Biol.* 20, 2774–2782
43. Yanagisawa, S. (2001) The transcriptional activation domain of the plant-specific Dof1 factor functions in plant, animal, and yeast cells. *Plant Cell Physiol.* 42, 813–822
44. Piskacek, S. *et al.* (2007) Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* 89, 756–768
45. Ma, J. *et al.* (1988) Yeast activators stimulate plant gene expression. *Nature* 334, 631–633
46. Erkina, T.Y. and Erkine, A.M. (2016) Nucleosome distortion as a possible mechanism of transcription activation domain function. *Epigenetics Chromatin* 9, 40
47. Nyanguile, O. *et al.* (1997) A nonnatural transcriptional coactivator. *Proc. Natl. Acad. Sci. U. S. A.* 94, 13402–13406
48. Saha, S. *et al.* (2003) RNA sequences that work as transcriptional activating regions. *Nucleic Acids Res.* 31, 1565–1570
49. Minter, A.R. *et al.* (2004) A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* 126, 10504–10505
50. Buhrlage, S.J. *et al.* (2009) Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors. *ACS Chem. Biol.* 4, 335–344
51. Erkina, T.Y. *et al.* (2010) Functional interplay between chromatin remodeling complexes RSC, SWI/SNF and ISWI in regulation of yeast heat shock genes. *Nucleic Acids Res.* 38, 1441–1449
52. Erkina, T.Y. *et al.* (2008) Different requirements of the SWI/SNF complex for robust nucleosome displacement at promoters of heat shock factor and Msn2- and Msn4-regulated heat shock genes. *Mol. Cell. Biol.* 28, 1207–1217
53. Scholes, N.S. and Weinzierl, R.O. (2016) Molecular dynamics of "Fuzzy" transcriptional activator–coactivator interactions. *PLoS Comput. Biol.* 12, e1004935
54. Hnisz, D. *et al.* (2017) A phase separation model for transcriptional control. *Cell* 169, 13–23
55. Staller, M.V. *et al.* (2018) A high-throughput mutational scan of an intrinsically disordered acidic transcriptional activation domain. *Cell Syst.* 6, 444–455 e446

56. Ravarani, C.N. *et al.* (2018) High-throughput discovery of functional disordered regions: investigation of transactivation domains. *Mol. Syst. Biol.* 14, e8190

57. Ma, J. and Ptashne, M. (1987) A new class of yeast transcriptional activators. *Cell* 51, 113–119

58. Abedi, M. *et al.* (2001) Transcriptional transactivation by selected short random peptides attached to lexA-GFP fusion proteins. *BMC Mol. Biol.* 2, 10

59. Pacheco, D. *et al.* (2018) Transcription activation domains of the yeast factors Met4 and Ino2: tandem activation domains with properties similar to the yeast Gcn4 activator. *Mol. Cell. Biol.* 38, e00038-18

60. Piskacek, M. *et al.* (2016) The 9aaTAD transactivation domains: from Gal4 to p53. *PLoS One* 11, e0162842

61. Csizmok, V. *et al.* (2018) Multivalent interactions with Fbw7 and Pin1 facilitate recognition of c-Jun by the SCF(Fbw7) ubiquitin ligase. *Structure* 26, 28–39 e22

62. Csizmok, V. *et al.* (2017) An allosteric conduit facilitates dynamic multisite substrate recognition by the SCF(Cdc4) ubiquitin ligase. *Nat. Commun.* 8, 13943