Unexpected Specificity within Dynamic Transcriptional Protein-Protein Complexes

*Matthew J. Henley[1,2], Brian M. Linhares[3], Brittany S. Morgan[1], Tomasz Cierpicki[2,3,4], Carol A. Fierke[5], Anna K. Mapp[1,2,6]**

[1]Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109
[2]Program in Chemical Biology, University of Michigan, Ann Arbor, MI 48109
[3]Department of Biophysics, University of Michigan, Ann Arbor, MI 48109
[4]Department of Pathology, University of Michigan, Ann Arbor, MI 48109
[5]Department of Chemistry, Texas A&M University, College Station, TX 77843
[6]Department of Chemistry, University of Michigan, Ann Arbor, MI 48109

**To whom correspondence should be addressed. Email: amapp@umich.edu*

Significance

Transcriptional activators represent a molecular recognition enigma; their function in transcription initiation requires selective engagement of coactivators yet the prevailing molecular recognition models propose this occurs via nonspecific intermolecular contacts. Here, mechanistic analysis of several related activator•coactivator complexes indicates a resolution to this enigma. In contrast to the expectations from nonspecific recognition models, even slight sequence changes in the activators cause these activator•coactivator complexes to undergo significant conformational redistribution, which is driven by specific intermolecular interactions and conformational changes in the coactivator itself. Our evidence reveals unappreciated *specific* molecular recognition mechanisms that underlie activator sequence variability, opening new questions about the relationship between recognition and function.

Dedication: We dedicate this work to Professor Laura L. Kiessling on the occasion of her 60th birthday.

Abstract

A key functional event in eukaryotic gene activation is the formation of dynamic protein-protein interaction networks between transcriptional activators and transcriptional coactivators. Seemingly incongruent with the tight regulation of transcription, many biochemical and biophysical studies suggest that activators use nonspecific hydrophobic and/or electrostatic interactions to bind to coactivators, with few if any specific contacts. In contrast, here a mechanistic dissection of a set of representative dynamic activator•coactivator complexes comprised of the ETV/PEA3 family of activators and the coactivator Med25 reveals a different molecular recognition model. The data demonstrate that small sequence variations within an activator family significantly redistribute the conformational ensemble of the complex while not affecting overall affinity, and distal residues within the activator—not often considered as contributing to binding—play a key role in mediating conformational redistribution. The ETV/PEA3•Med25 ensembles are directed by specific contacts between the disordered activator and the Med25 interface, and this specificity is facilitated by structural shifts of the coactivator binding surface. Taken together, this highlights the critical role coactivator plasticity plays in recognition of disordered activators, and indicates that molecular recognition models of disordered proteins must consider the ability of the binding partners to mediate specificity.

Introduction

Protein-protein interactions (PPIs) formed between transcriptional activators and coactivators play a critical role in gene expression; these PPIs underpin co-localization of transcriptional machinery components and stimulate transcription initiation.(1–8) A prevailing view of activator-coactivator PPIs is that they are largely non-specific and, further, that the selectivity necessary for appropriate gene expression comes from other sources such as activator-DNA interactions and/or co-localization.(3, 9–14) Indeed, there is considerable data suggesting activator•coactivator complexes form via almost entirely nonspecific intermolecular interactions, from early experiments demonstrating that a wide range of natural and non-natural amphipathic molecules interact with coactivators to more recent structural studies indicating no fixed activator•coactivator binding mode.(9–11, 15–18)

Nonspecific recognition models, while attractive in their simplicity, are inconsistent with the critical functional role that individual activator•coactivator PPIs play in gene expression. There are several examples of transcriptional activators that depend on interactions with specific activator binding domains (ABDs) of coactivators for function,(19–23) e.g., the SREBP family of activators require the KIX ABD of the coactivator ARC105 to regulate fatty acid homeostasis(21) even though other coactivators such as CBP/p300 have structurally similar KIX motifs. Further, the biophysical studies that investigate how ABDs recognize diverse activators most often utilize qualitative equilibrium approaches(9–11) that are blind to critical mechanistic information(24–26) due to equilibrium averaging. It is therefore an open question whether

there are other molecular recognition mechanisms at play; this would account for the diversity of functional activator sequences as well as the observed selectivity of activators *in vivo*. Because activator•coactivator complexes often represent promising therapeutic targets, developing a more detailed understanding of the molecular recognition mechanisms of these crucial PPIs is also essential for the development of small molecule modulators.(4, 27–29)

Here, we take a critical look at activator•coactivator recognition by mechanistically dissecting a representative set of dynamic complexes formed between the ABD of Mediator subunit Med25 and the amphipathic transcriptional activation domains (TADs) of the ETV/PEA3 family of Ets transcriptional activators (ETV1, ETV4, and ETV5).(23, 30–32) Previous biophysical studies indicated that the interaction of Med25 with family member ETV5 appears to be a prototypical nonspecific TAD•ABD complex: it occurs over a shallow surface, is driven by electrostatic and hydrophobic interactions, and forms a dynamic complex that is recalcitrant to structure determination.(30, 32) We utilize a mechanistic and structural approach that combines quantitative data regarding activator•coactivator conformational states obtained via transient kinetic analysis with structural information obtained through mutagenesis and NMR spectroscopy. Our data reveal that the conformational ensembles of ETV/PEA3•Med25 PPIs are strikingly sensitive to slight changes in TAD sequence, despite being dynamic complexes with several well-populated conformational sub-states at equilibrium. Furthermore, the mechanisms by which conformational sensitivity arises are complex and intertwined, involving the ability of ordered and disordered regions of the TAD to participate in finite

sets of alternative interactions with the Med25 interface, as well as conformational changes in Med25 that remodel the TAD binding site. Together, these results reveal an unappreciated degree of specificity in the formation of activator•coactivator complexes that is in direct contrast to the prevailing nonspecific recognition models of these essential PPIs.(3, 11)

Results

ETV/PEA3•Med25 PPIs as a Model System for Dynamic TAD•ABD Interactions

ETV/PEA3•Med25 interactions represent an ideal system to study dynamic TAD•ABD interactions for several reasons. First, previous studies indicated that interaction of ETV/PEA3 family member ETV5 with the Med25 ABD is typical of a dynamic TAD•ABD complex(30, 32); the binding surface is shallow, both acidic and hydrophobic amino acids of the TAD determine affinity and activity, and multiple bound conformational states were detected by both NMR and kinetic analyses. Second, our previous studies of the ETV5•Med25 complex showed that the bound conformational ensemble is directly accessible by transient kinetic analysis.(32) The models of TAD•ABD molecular recognition can thus be dissected in this system without the loss of critical conformational information to equilibrium averaging. Third, the ETV/PEA3 family of transcription factors serves as an excellent natural system to test the relationship between TAD sequence and recognition; they contain almost identical arrangements of acidic and hydrophobic residues across the TAD sequence, especially within the helical binding region that undergoes coupled folding and binding with Med25,(30) but the identity of specific residues varies slightly (Fig. 1A). This system can therefore be used to test whether

TAD•Med25 interactions are truly nonspecific and insensitive to changes in the TAD sequence (i.e., if these PPIs form in a “sequence-independent” manner(11)), or if these interactions are affected by TAD sequence changes and thus have a degree of specificity to formation. Finally, the Med25 ABD is ligandable by small molecules(32), and therefore conclusions from mechanistic studies can be directly applied to guide and assess optimization of small molecule modulators of TAD•Med25 complex formation.

Small Sequence Differences Between ETV/PEA3 Family Members Lead to Conformationally Distinct PPIs with Med25

We performed stopped-flow fluorescence transient kinetic experiments to directly measure the effects of TAD sequence differences on the conformational ensemble of individual ETV/PEA3•Med25 complexes, using TADs synthesized with the solvatochromic fluorophore 4-DMN conjugated to the *N*-terminus.(33) We previously demonstrated with this approach that the ETV5•Med25 complex forms in a minimal three step linear mechanism(32): after an initial rapid association event that mostly occurs in the instrument dead-time (~2-4 ms), the complex undergoes two sequential conformational changes (Fig 1B). Application of the same experimental conditions to ETV1 and ETV4 indicated that the kinetic binding mechanism is conserved; for all ETV/PEA3 TADs a rapid binding step followed by two conformational change steps was observed. Each of these individual steps occurred with similar exchange rate constants (k_{ex}; the sum of the forward and reverse rate constants) for each ETV/PEA3•Med25 complex (Fig. 1B,C); this suggests that the steps represent analogous conformational transitions in each complex. In addition, the equilibrium binding affinity between

ETV/PEA3 TADs varied less than two-fold (0.7–1.2 μM, Fig. 1C), consistent with the expectation from nonspecific models that minor substitutions in the TAD will not affect the overall stability of the activator•coactivator complex.

Despite a conserved binding mechanism and similar overall affinities, calculation of equilibrium conformational populations from the kinetic data revealed clear differences between the engagement modes of ETV/PEA3 family members (Fig. 1C,D; for raw data and detailed kinetic analysis, see SI Discussion of Kinetic Analysis). While the populations of analogous conformational states of the ETV1•Med25 and ETV5•Med25 complexes were essentially identical, the ETV4•Med25 complex populated the three analogous conformations in a unique manner (Fig 1C,D). Critically, this shift in conformational ensemble does not correlate with predicted structural propensity differences between the ETV/PEA3 TADs (SI Fig. S1), which suggests that variable residues between ETV1, ETV4, and ETV5 alter the TAD•Med25 conformational ensemble via intermolecular interactions made in the bound state.

We next sought to confirm whether there are structural differences between these complexes using NMR spectroscopy. Side-chain methyl ^{1}H,^{13}C-HSQC experiments were used as a primary method to enable direct detection of effects on both surface and buried residues of Med25. Comparative analysis of Med25 ^{1}H,^{13}C-HSQC spectra bound to different ETV/PEA3 family members was consistent with the expected engagement differences between the ETV/PEA3 TADs: spectra of ETV1- and ETV4-bound Med25 exhibited several large differences in chemical shift perturbation (CSP) patterns, whereas

the spectra of Med25 bound to ETV1 and ETV5 were essentially indistinguishable (Fig. 1E,F; full spectral overlay shown in SI Fig. S19). Inspection of the CSP data plotted on the structure of Med25 indicated that all ETV/PEA3 family members bind to a previously identified(30–32) core binding site formed between the central β-barrel and the *C*-terminal α3 helix (Fig. 1G), but ETV1/ETV5 and ETV4 produce unique perturbation patterns in the binding surface (Fig. 1G, cyan circles). In addition, several resonances representing buried and/or allosteric residues displayed significant CSP differences between the ETV1- and ETV4-bound complexes, suggesting that the conformation of the Med25 ABD may also be different between these complexes (Fig. 1F, starred). Together, the HSQC data supports the conclusion from kinetics experiments that ETV4 has a unique Med25 engagement mode as compared to ETV1 and ETV5.

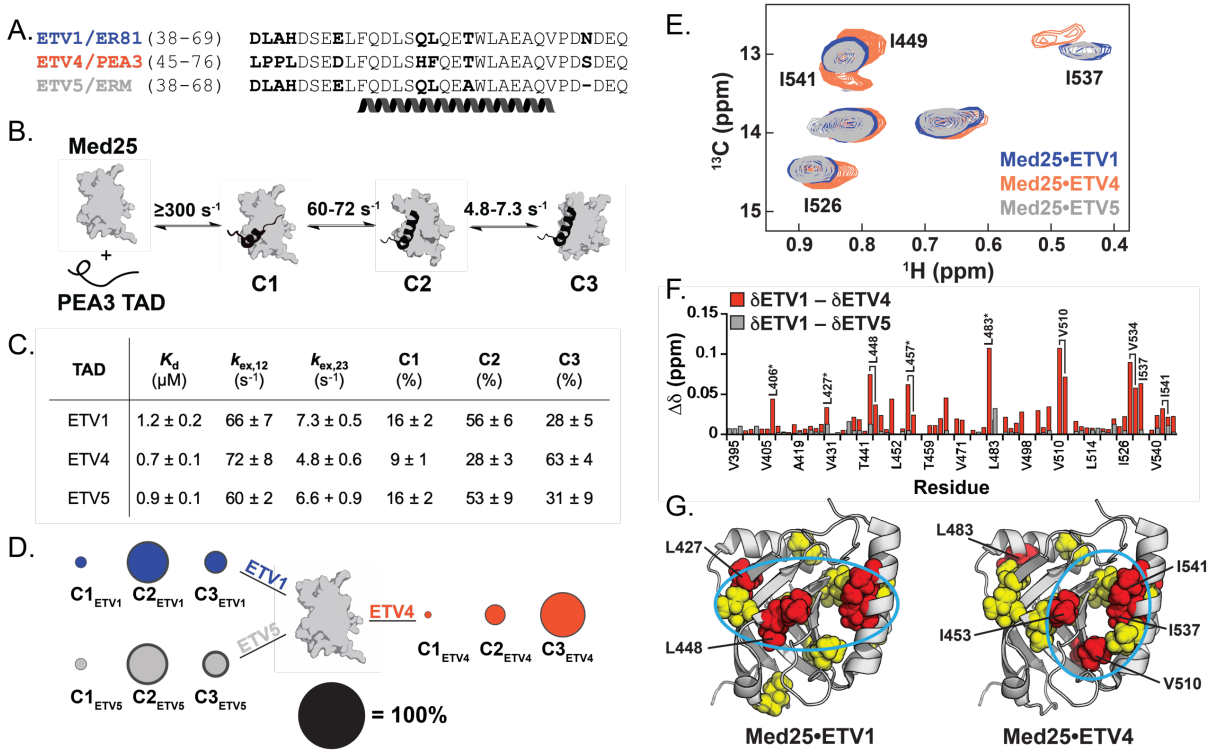


Figure 1. ETV/PEA3 activators differentially engage with Med25. A) Alignment of ETV/PEA3 family activation domains. The helix denotes the residues that undergo coupled folding and binding with Med25, as determined by NMR chemical shift analysis.*(30)* B) Mechanism of binding of ETV/PEA3 activators to Med25, determined here for ETV1 and ETV4, and previously for ETV5.*(32)* The range of exchange rates between analogous steps for ETV/PEA3 TADs are shown. C) Table of relevant binding parameters for each ETV/PEA3 TAD, including the equilibrium affinity, exchange rates between C1 and C2 ($k_{ex,12}$) and between C2 and C3 ($k_{ex,23}$), and equilibrium populations of each state. All values represent the average of 3-4 biological replicates, and the error is the standard deviation. D) Equilibrium populations of the three ETV/PEA3•Med25 conformations scaled relative to the diameter of the black circle. Standard deviations of the values are shown as the dark gray outer circle. E) Overlay of the Ile Cδ region of the ^{1}H,^{13}C-HSQC of Med25 in complex with 1.1 equivalents unlabeled ETV1 (blue), ETV4 (orange), or ETV5 (grey). Peaks that have chemical shift differences between complexes are labeled. Note: I449 and I541 form a single overlapped peak for the ETV1 and ETV5 complexes. Full spectra are in the supporting information. F) Chemical shift differences of Med25 methyl resonances between ETV1 and ETV4 (orange) and ETV1 and ETV5 (grey). G) Chemical shift perturbations induced by binding of ETV/PEA3 activators plotted on the structure of Med25 (PDB ID 2XNF)*(34)*. Yellow = 0.040–0.080 ppm, red > 0.080 ppm. Cyan circles highlight general distinctions in perturbation patterns. Several residues with chemical shift differences >0.030 ppm between the ETV1•Med25 (or ETV5•Med25) and ETV4•Med25 complexes are labeled.

Ordered and Disordered Regions of the ETV/PEA3 TADs Dictate Conformational Differences between ETV/PEA3•Med25 PPIs

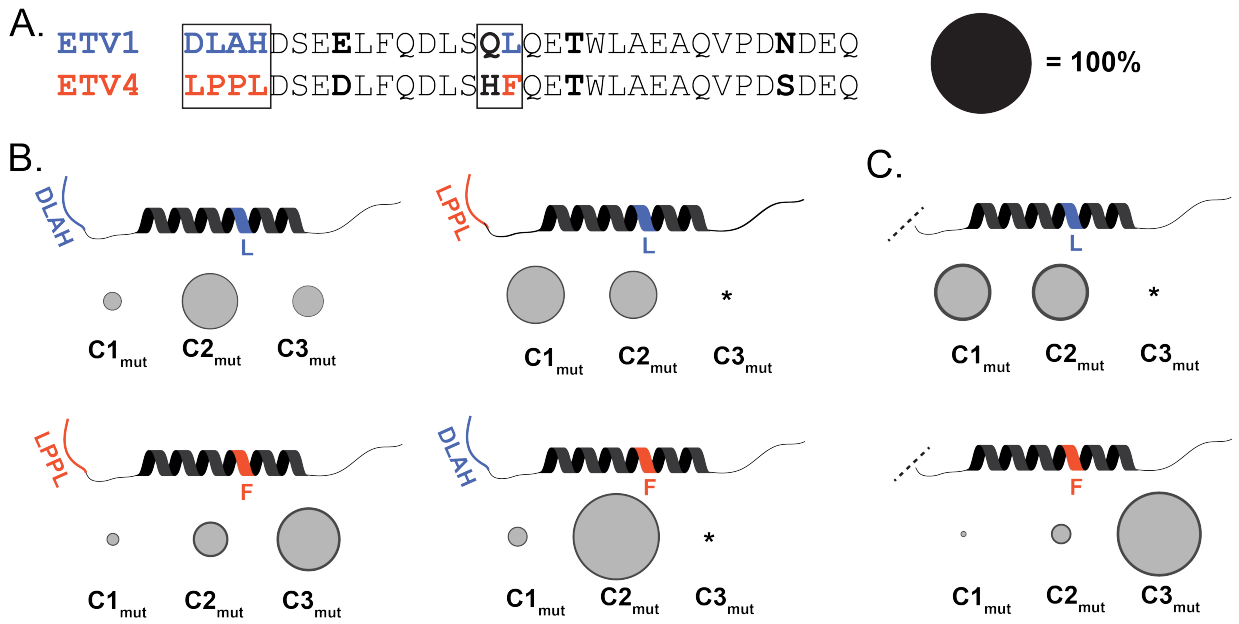
We next used a mutagenesis approach to examine the precise differences between the activator sequences that bias ETV/PEA3•Med25 PPIs towards distinct conformational sub-states. This effort focused on residues that are conserved between ETV1 and ETV5, but not ETV4. Two regions of interest were evident (Fig. 2A, boxed): 1) a two amino acid 'variable motif' in the helical binding region consisting of a polar residue followed by a hydrophobic residue (QL in ETV1/ETV5 and HF in ETV4), and 2) the four amino acid *N*-terminus of the TAD (DLAH in ETV1/ETV5 and LPPL in ETV4), a region predicted to be entirely disordered for all ETV/PEA3 TADs (SI Fig. S1 and Ref. 30). A small library of mutant TADs based on the variations of the ETV4 scaffold in the two regions of interest—either the DLAH or LPPL *N*-terminus with Leu or Phe residues in the helical binding region—was synthesized and assessed in stopped-flow kinetic assays (Fig. 2B; values for all variants in SI Tables S2 and S3). The polar residue (Gln/His) in the variable motif was also tested but had no effect on conformational populations. Importantly, the equilibrium binding affinity varied only slightly across the mutants (Fig. 2D).

Consistent with the hypothesis that one or both of the variable regions dictate the conformational differences between native ETV/PEA3•Med25 complexes, ETV/PEA3 variants with 'native' combinations of the *N*-terminus and the hydrophobic residue of the 'variable motif' exhibited similar conformational sub-state populations as compared to the native TADs (Fig. 2B, left). That is, combinations with DLAH at the *N*-terminus and Leu in the variable motif (DLAH/L) populated three observable conformational states in a similar

manner to ETV1 and ETV5 (Fig. 2B, top left), and, analogously, combinations with LPPL at the *N*-terminus and Phe in the variable motif (LPPL/F) populated three sub-states comparably to ETV4 (Fig. 2B, bottom left).

Conversely, when nonnative combinations of the *N*-terminus and variable motif were tested, unique conformational behavior was observed (Fig. 2B, right). Only two bound conformations of DLAH/F variants were detected in kinetics experiments, and displayed exchange kinetics similar to the C1–C2 transitions of the native complexes (Fig. 2B, top right). Calculation of conformational populations indicated that the second conformation had a higher overall population (82 ± 3%) than the C2 conformations of the ETV1 (56 ± 6%), ETV5 (53 ± 9%), or ETV4 complexes (28 ± 3%). Similarly, two bound conformations with C1–C2-like exchange rates were detected with LPPL/L variants (Fig 2B, bottom right), but the initial bound sub-state was preferentially populated. We note that the C3 conformation of these complexes may be undetectable due to a low population (≤10%) or an increase in the exchange kinetics between C2–C3; however, neither of these possibilities are inconsistent with the conclusion that these mutant complexes are conformationally distinct (see SI *Discussion of Kinetic Data Analysis* for further details). Together, these results indicate a model where *both* the hydrophobic residue in the variable motif and the disordered *N*-terminus dictate the conformational differences between ETV/PEA3•Med25 complexes. The latter result is particularly striking, as disordered regions of the TAD are often removed/ignored in biophysical and structural studies because they typically do not contribute to overall affinity.(9–11, 25, 35–40)

To obtain further evidence for the unexpected role of the disordered *N*-terminus on the conformational behavior of ETV/PEA3•Med25 PPIs, we also tested the effects of removing the four variable *N*-terminal residues (ΔNt) of ETV1 and ETV4 in kinetics experiments. The resulting variants ETV1$^{\Delta Nt}$ and ETV4$^{\Delta Nt}$ displayed differential changes in conformational behavior from the parent TADs (Fig. 2C), in addition to a slight (~1.7-fold) gain in affinity for both variants (Fig. 2D). Removal of the ETV1 N-terminus resulted in significant conformational redistribution; kinetic analysis indicated that the ETV1$^{\Delta Nt}$•Med25 complex exchanged between two equally populated conformational sub-states on a similar timescale to the C1–C2 transition of the parent ETV1•Med25 complex. On the other hand, the ETV4$^{\Delta Nt}$•Med25 complex populated three conformational sub-states in an analogous manner to the parent ETV4•Med25 complex. Together, these results therefore support a direct role of the *N*-terminal residues of the ETV1/ETV5 TADs, but not the ETV4 TAD, in dictating the conformational behavior of the native TAD•Med25 complexes.



D.

Variant	DLAH/L	DLAH/F	LPPL/L	LPPL/F	ETV1$^{\Delta Nt}$	ETV4$^{\Delta Nt}$
K_d (μM)	1.1 ± 0.2	0.9 ± 0.1	1.9 ± 0.2	0.7 ± 0.1	0.7 ± 0.1	0.4 ± 0.1

Figure 2. Variable residues in the disordered *N*-terminus and the helical binding region mediate differences in ETV/PEA3•Med25 conformational behavior. A) Alignment of ETV1 and ETV4 activators with regions that were selected for mutational analysis boxed. Regions/residues that affected the conformational ensemble are color coded to ETV1 (blue) or ETV4 (orange). Effects of the Gln/His residues in the variable motif were also tested but did not affect conformational populations and are thus omitted in B. and C. for clarity. Populations of conformational states in B. and C. are scaled relative to the diameter of the black circle. B) Results from kinetics experiments of mutant TADs, for native (left) and non-native (right) combinations of variable *N*-termini and helical binding regions. Variants were made based on the ETV4 sequence. The data shown is the average across all the variants tested from each group, with the error (dark grey outer circle) representing the standard deviation. C) Results from kinetics experiments with ETV1$^{\Delta Nt}$ (left) and ETV4$^{\Delta Nt}$ (right). D) Average equilibrium K_d values of variants tested. *Conformer was undetectable in kinetics experiments (see SI *Discussion of Kinetic Data Analysis* for further details).

Variable Regions of the ETV/PEA3 TADs Differentially Engage with the Med25 Surface

An NMR strategy to probe the structural basis by which the two key regions in the ETV/PEA3 TADs modulate the bound ETV/PEA3•Med25 conformations was next

pursued. We utilized a conservative mutagenesis approach, where minimally perturbing mutations were individually introduced into unlabeled TADs, and then CSP analysis of $^{1}H,^{13}C$-HSQC spectra of Med25 bound to the native or mutated TAD was performed to identify Med25 methyl groups affected by the mutation. Analysis of CSP differences in the mutant TAD•Med25 HSQC spectrum can therefore detect Med25 residues in direct proximity to the mutated site in the complex and also has the potential to reveal allosteric connections if the effects of the mutation are propagated from the interaction site.(41) Furthermore, this strategy avoids the significant experimental challenge associated with NMR analysis of Med25-bound ETV/PEA3 TADs, where a significant fraction of peaks are too broad to detect due to conformational exchange.(30) Here, conservative mutations were introduced into the two key variable regions of the ETV1 and ETV4 TADs to detect differences in engagement modes that could explain the effects of these regions on ETV/PEA3•Med25 conformational states.

Mutations were first made within the variable motif of the helical binding region by swapping the variable polar residue between ETV1 and ETV4 to form $ETV1^{Q52H}$ and $ETV4^{H59Q}$ (note: residue numbers for ETV4 are shifted by +7 compared to ETV1 and ETV5), based on the observation that this change did not affect the populations of conformational states in kinetics experiments (SI Table S3). Indeed, the $^{1}H,^{13}C$-HSQC spectra of the mutant ETV/PEA3•Med25 complexes were almost identical to those of the native complexes except for single shifts in unique methyl peaks of Ile541 (Fig. 3A): the $ETV4^{H59Q}$ variant produced a large shift (0.032 ppm) in the Med25 Ile541δ peak, whereas the analogous $ETV1^{Q52H}$ variant produced a smaller perturbation (0.015 ppm) in the

Ile541γ peak. These highly localized shifts are consistent with the mutations causing proximity-based perturbations near the mutation site in the complex. In addition, the fact that the individual mutations perturb unique methyl groups originating from the same residue suggest that the native ETV1 QL and ETV4 HF motifs are engaged in unique interactions in a similar interface.

Next, we introduced mutations into the disordered *N*-terminus. Specific sites for mutations in this region were not immediately obvious from previous data, so several point mutations were made; Leu to Val mutations ETV1^{L39V} and ETV4^{L48V} sufficed to produce measurable differences in bound ^{1}H,^{13}C-HSQC spectra compared to the native TADs, without significantly altering the overall spectra (Fig. 3B, SI Figs. S22 and S23). In contrast to mutations in the helical binding region, ETV1^{L39V} and ETV4^{L48V} affected a larger overall site on the Med25 structure, indicating possible indirect effects stemming from these mutations. Indeed, both of these *N*-terminal mutants altered an overlapping sub-set of residues in the core binding site, which would be occupied by the TAD helical binding region. Nonetheless, evidence for a direct interaction of the ETV1 *N*-terminus was also apparent from the ETV1^{L39V}•Med25 spectrum; several shifts from the native ETV1•Med25 spectrum were observed in a cluster of residues in a distal site involving Val405, Leu427, Thr476, and Leu483. Conversely, this site was unaffected by the ETV4^{L48V} variant, suggesting this interaction is made only by the ETV1 *N*-terminus. Consistent with this representing a functional interaction of the ETV1 *N*-terminus, our kinetics data demonstrated that removal of the ETV1 *N*-terminus significantly alters the ETV1•Med25

conformational ensemble, whereas removal of the ETV4 *N*-terminus only slightly affected the ETV4•Med25 ensemble (Fig. 2C).

To further test this model, we synthesized ETV1 and ETV4 TADs that were selectively ^{15}N-labeled at single Leu residues and analyzed CSPs in ^{1}H,^{15}N-HSQC spectra upon binding of unlabeled Med25. Analysis of Leu residues in the helical binding region was attempted by this method; however, in almost all cases these peaks were too broad to detect when the TADs were bound to unlabeled Med25. In contrast, peaks for Leu residues in the *N*-terminus remained relatively sharp upon binding to Med25, likely due to these regions retaining more structural disorder in the complex.(30) Comparison of CSPs of Leu residues in the *N*-termini of ETV1 and ETV4 were consistent with the proposed differential interaction with Med25: Leu39 in the ETV1 *N*-terminus underwent a large (~0.2 ppm) shift whereas Leu48 in the ETV4 *N*-terminus shifted only slightly (Fig. 3C) upon addition of unlabeled Med25. Interestingly, we also observed a minor peak in spectra of bound ETV1 Leu39 that was slightly shifted from the unbound position, perhaps representing one of the lowly-populated ETV1•Med25 conformations (Fig. 1D) where the *N*-terminus is weakly bound to the Med25 surface.

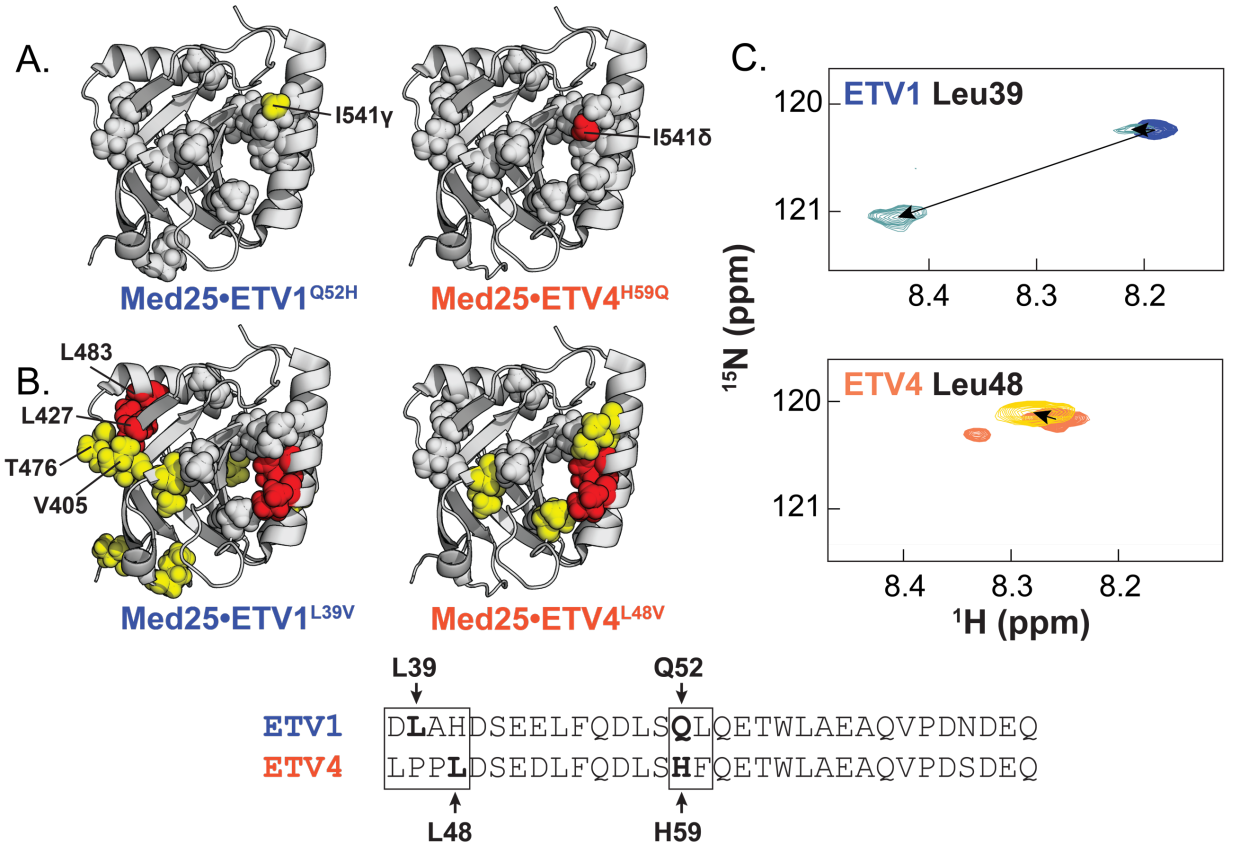


Figure 2. ETV/PEA3 variable regions engage in unique interactions with the Med25 surface. Effects of conservative mutations in the A) helical binding region and B) *N*-termini are plotted on the structure of Med25. Yellow = 0.015 – 0.030 ppm. Red ≥ 0.030 ppm. Residues discussed in the text are labeled. Grey spheres denote residues that undergo identical perturbations in both parent and mutant complexes. Residues chosen for mutation are bolded and labeled in the alignment. C) Chemical shift perturbations of 150 μM ETV1 (above) and ETV4 (below) TADs in the absence (blue and orange, respectively) and presence (light blue and maroon, respectively) of 280 μM unlabeled Med25. TADs were selectively ^{15}N labeled at the positions noted. Small secondary peaks in free ETV4 spectra were observed and likely arose from isomerization of the two tandem Pro residues in the *N*-terminal region.

ETV/PEA3•Med25 Conformational Changes Involve Shifts in Med25 Structure

Altogether, these data support a model where the conformational differences between ETV/PEA3•Med25 complexes are caused by the ETV/PEA3 variable regions engaging with the Med25 surface in unique and sequence-dependent manners. While NMR and mutagenesis data revealed that the variable *N*-terminus affects ETV/PEA3•Med25

conformation via differential engagement with the Med25 surface, the mechanism by which bound conformational behavior is affected by the variable motif in the helical binding region is unclear. Interestingly, NMR analysis demonstrated that the ETV1 and ETV4 variable motifs localize to a similar region (Fig. 3A), which indicates that conformational differences caused by this motif originate from distinct interactions in the same site. We thus reasoned that this would likely involve remodeling of the Med25 ABD. To test this hypothesis, we examined changes in the bound TAD•Med25 ^{1}H,^{13}C-HSQC spectra produced by a point mutation in the variable motif that significantly redistributes the populations of conformational states. The ETV4^{F60L} mutation was selected because this small change in residue identity caused a drastic conformational redistribution (Fig. 2B, compare bottom left and top right) to favor the initial bound sub-state (C1).

Comparison of the ^{1}H,^{13}C-HSQC spectra of ETV4•Med25 and ETV4^{F60L}•Med25 revealed several Med25 peaks that are significantly perturbed when bound to native ETV4 move towards the unbound position in the ETV4^{F60L}•Med25 spectrum (Fig 4A). Significantly, the ETV4^{F60L} variant elicited weaker CSPs around the core binding site than native ETV4 (Fig. 4B), suggesting that several of the large CSPs in this region are tied to the conformational changes. Furthermore, this behavior was observed for peaks representing residues that are buried or in allosteric regions of the protein, including the β-barrel core, the interface between the β-barrel and the allosteric α2 helix, and the interface between the *C*-terminal α3 helix and the allosteric α1 helix (Fig. 4B). These data are therefore consistent with a direct role for ABD conformational plasticity in molecular recognition,

which likely enables specific recognition by revealing new topology in the core binding site.

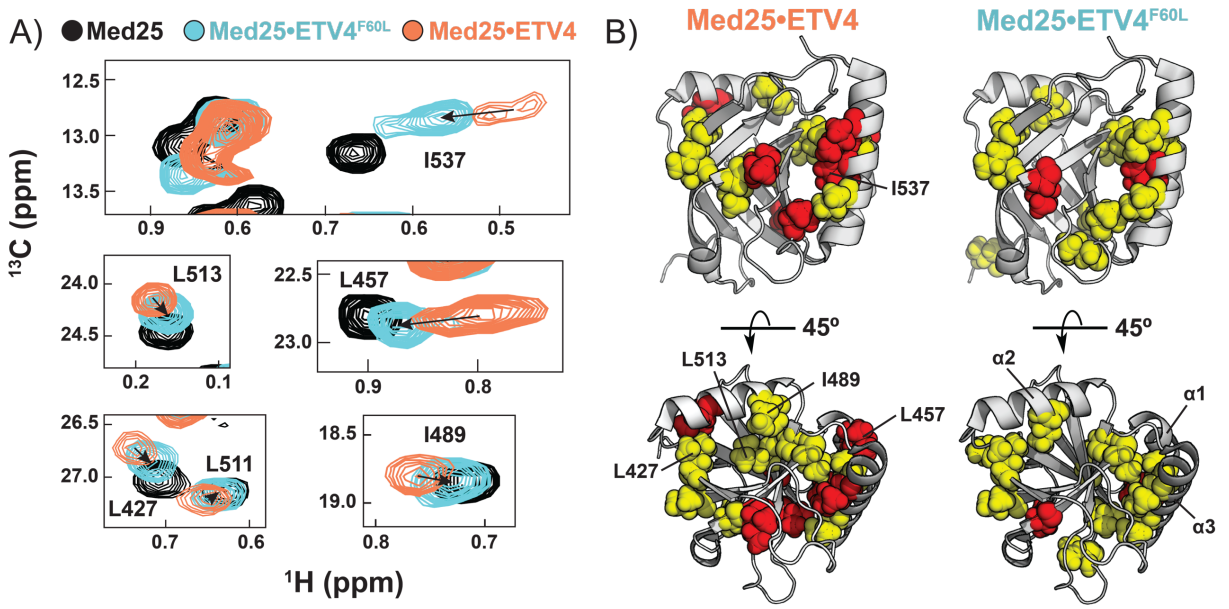


Figure 4. Structural shifts in the Med25 ABD accompany the conformational changes of ETV/PEA3•Med25 complexes. A) Examples of Med25 resonances undergoing shifts toward the unbound position (black) upon ETV4^{F60L} (cyan) mutation. B) Chemical Shift Perturbations from binding of ETV4 (left) and ETV4^{F60L} (right). Residues shown in A. are labeled on the structures. Yellow = 0.040–0.080 ppm, red > 0.080 ppm.

Discussion

The exceptional sequence variability of functional TADs—characterized only by a general preponderance of acidic and hydrophobic amino acids—has remained a molecular recognition enigma over the past several decades.(1–5, 10, 12–14) There have been several recognition models advanced to account for the large variety of functional TADs, most of which propose that TAD•coactivator recognition occurs via nonspecific intermolecular interactions.(3, 9–12) The major driving force of association in these

models is thus the stochastic burial of hydrophobic sidechains rather than the formation of defined intermolecular contacts typical for well-structured PPIs. However, there is limited direct biophysical evidence for nonspecific mechanisms, and the available biophysical data relies almost entirely on a) the observation of similar CSP patterns of unique TAD sequences binding to the same ABD,(11) and b) NMR paramagnetic relaxation enhancement (PRE) patterns showing multiple bound orientations of the TAD on the ABD surface.(9–11) Alternative explanations also exist for both of these observations: similar CSP trajectories are expected if the ABD undergoes a conserved conformational change upon binding,(42–44) and significant PREs can be observed from lowly populated (<10%) sub-states that place the paramagnetic spin label in close proximity to the interacting partner.(45, 46) The degree to which activator•coactivator recognition is truly nonspecific in such examples is therefore unclear. Unfortunately, this widely accepted view of activator•coactivator recognition also represents a primary reason why these PPIs have been traditionally considered “untargetable”.(29)

Here, we scrutinized these recognition models by subjecting the dynamic PPIs formed between the ETV/PEA3 family of TADs and their binding partner Med25 to detailed mechanistic and structural dissection. We found that these interactions exhibited a striking degree of conformational sensitivity to small changes in TAD sequence, which is inconsistent with PPIs driven by nonspecific intermolecular interactions. Instead, both ordered and disordered regions of ETV/PEA3 TADs have the capacity to engage in a finite set of alternative interactions with the Med25 surface. This recognition mechanism is further enabled by an underappreciated role of ABD plasticity in molecular recognition:

in contrast to the shallow and largely featureless ETV/PEA3 binding surface presented by Med25 in the unbound state, the Med25 ABD undergoes significant remodeling upon complex formation and thus likely plays a direct role in enabling different interaction modes.(47)

Taken together, these data reveal an alternative mechanism of TAD•ABD recognition where specificity underlies the sequence variability of functional TADs. In contrast to nonspecific mechanisms, the model of TAD•ABD molecular recognition outlined here highlights a direct path to targeting strategies for small molecule therapeutics or probes of activator•coactivator complexes. Specifically, the role of ABD plasticity in recognition suggests that development of molecules that stabilize specific ABD conformational states may be a more effective targeting strategy than directly targeting the topologically challenging TAD•ABD interface. We previously obtained allosteric modulators of Med25 and other ABDs by covalent targeting of dynamic structural elements,(32, 48) which supports the idea that ABD plasticity can be effectively exploited by small molecules.

A fundamental biological question that emerges from our study is how molecular recognition mechanisms affect function. In general, it is still exceptionally challenging to relate PPI function and affinity; in processes where dynamic PPIs serve as critical functional events, such as transcription and proteostasis, there is often little or no correlation between affinity and functional activity.(49, 50) Several factors play into this observation, such as subcellular localization and concentration, but a potentially significant factor is the mechanism by which the complex forms. For example, in our

current study we observed that the lifetimes of individual ETV/PEA3•Med25 conformations varied up to two orders of magnitude (SI Table S2), therefore shifting the conformational ensemble towards longer-lived sub-states could have significant functional outcomes without necessitating changes in affinity.(51) This work therefore highlights a critical need to understand the relationship between the function of biomolecular interactions and the mechanisms by which they are formed.

Methods

Protein Expression and Purification

Med25 AcID was expressed and purified from a pET21b-Med25(394-543)-His_6 plasmid from *E. coli* BL21 (DE3) cells as described previously (full details in the Supporting Information).(32) Uniformly ^{13}C,^{15}N labeled Med25 for NMR experiments was expressed identically except using M9 minimal media supplemented with 1 g/L $^{15}NH_4Cl$, 2 g/L ^{13}C-$_D$-glucose, and 0.5% ^{13}C,^{15}N-labelled Bioexpress media. Protein identity was confirmed by mass spectrometry (Agilent Q-TOF).

Peptide Synthesis

The peptides used in this study were prepared using standard Fmoc solid-phase peptide synthesis on a Liberty Blue Microwave Peptide Synthesizer (CEM). ^{15}N-Leu labeled peptides were synthesized using Fmoc-^{15}N-Leu in place of unlabeled Fmoc-Leu at the specified positions. Details of synthesis and peptide characterization data (HPLC) are included in the Supporting Information.

Stopped-flow kinetics

Stopped-flow kinetic assays were performed using a Kintek SF-2001 stopped flow instrument equipped with a 100-W Xe arc lamp in two-syringe mode. All experiments were completed at 10 ºC in stopped-flow buffer (10 mM sodium phosphate, 100 mM NaCl, 2% DMSO, 1% glycerol, 0.001% NP-40, pH 6.8). All concentrations reported are after mixing. The 4-DMN fluorophore was excited at 440 nm, and fluorescence intensity was measured at wavelengths >510 nm using a long-pass filter (Corion). Further details, including an explanation of how reported values were determined, are in the Supporting Information.

NMR Spectroscopy

Constant time ^{1}H,^{13}C-HSQC experiments were performed with uniformly ^{13}C,^{15}N-labeled Med25 in NMR buffer (20 mM sodium phosphate pH 6.5, 150 mM NaCl, 3 mM DTT, 10% D_2O, and 2% DMSO) on a Bruker 600 MHz instrument equipped with a cryoprobe. HSQC experiments were processed in NMRPipe(52) and visualized with NMRFAM-Sparky.(53) Chemical shift perturbation analyses were performed on samples with 1.1 equivalents of unlabeled binding partner, which results in ≥96% bound Med25 based on measured K_d values. Assignments of side-chain methyl resonances of free Med25 were achieved through 3D H(CCCO)NH and (H)CC(CO)NH TOCSY experiments (23 ms TOCSY mixing time) performed with a sample of 600 μM ^{13}C,^{15}N Med25 on a Bruker 800 MHz instrument equipped with a cryoprobe. Further details about assignment and data analysis can be found in the Supporting Information.

References

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

2. I. A. Hope, S. Mahadevan, K. Struhl, Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333**, 635–640 (1988).

3. P. B. Sigler, Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).

4. A. K. Mapp, A. Z. Ansari, A TAD Further: Exogenous Control of Gene Activation. *ACS Chem. Biol.* **2**, 62–75 (2007).

5. S. Piskacek, *et al.*, Nine-amino-acid transactivation domain: Establishment and prediction utilities. *Genomics* **89**, 756–768 (2007).

6. M. Krishnamurthy, *et al.*, Caught in the Act: Covalent Cross-Linking Captures Activator–Coactivator Interactions *in Vivo*. *ACS Chem. Biol.* **6**, 1321–1326 (2011).

7. A. Boija, *et al.*, Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).

8. B. R. Sabari, *et al.*, Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

9. P. S. Brzovic, *et al.*, The Acidic Transcription Activator Gcn4 Binds the Mediator Subunit Gal11/Med15 Using a Simple Protein Interface Forming a Fuzzy Complex. *Mol. Cell* **44**, 942–953 (2011).

10. L. Warfield, L. M. Tuttle, D. Pacheco, R. E. Klevit, S. Hahn, A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci.* **111**, E3506–E3513 (2014).

11. L. M. Tuttle, *et al.*, Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep.* **22**, 3251–3264 (2018).

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

13. C. D. Arnold, *et al.*, A high-throughput method to identify trans-activation domains within transcription factor sequences. *EMBO J.* **37** (2018).

14. M. V. Staller, *et al.*, A High-Throughput Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. *Cell Syst.* **6**, 444-455.e6 (2018).

15. A. R. Minter, B. B. Brennan, A. K. Mapp, A Small Molecule Transcriptional Activation Domain. *J. Am. Chem. Soc.* **126**, 10504–10505 (2004).

16. S. P. Rowe, A. K. Mapp, Assessing the permissiveness of transcriptional activator binding sites. *Biopolymers* **89**, 578–581 (2008).

17. S. J. Buhrlage, *et al.*, Amphipathic Small Molecules Mimic the Binding Mode and Function of Endogenous Transcription Factors. *ACS Chem. Biol.* **4**, 335–344 (2009).

18. C. Y. Majmudar, *et al.*, Sekikaic Acid and Lobaric Acid Target a Dynamic Interface of the Coactivator CBP/p300. *Angew. Chem. Int. Ed.* **51**, 11258–11262 (2012).

19. J. C. Chrivia, *et al.*, Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365**, 855–859 (1993).

20. J. L. Ruas, L. Poellinger, T. Pereira, Functional Analysis of Hypoxia-inducible Factor-1α-mediated Transactivation: IDENTIFICATION OF AMINO ACID RESIDUES CRITICAL FOR TRANSCRIPTIONAL ACTIVATION AND/OR INTERACTION WITH CREB-BINDING PROTEIN. *J. Biol. Chem.* **277**, 38723–38730 (2002).

21. F. Yang, *et al.*, An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442**, 700–704 (2006).

22. D. Sela, *et al.*, Role for Human Mediator Subunit MED25 in Recruitment of Mediator to Promoters by Endoplasmic Reticulum Stress-responsive Transcription Factor ATF6α. *J. Biol. Chem.* **288**, 26179–26187 (2013).

23. A. Verger, *et al.*, The Mediator complex subunit MED25 is targeted by the N-terminal transactivation domain of the PEA3 group members. *Nucleic Acids Res.* **41**, 4847–4859 (2013).

24. D. D. Boehr, D. McElheny, H. J. Dyson, P. E. Wright, The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis. *Science* **313**, 1638–1642 (2006).

25. K. Sugase, H. J. Dyson, P. E. Wright, Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025 (2007).

26. P. Vallurupalli, D. F. Hansen, L. E. Kay, Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc. Natl. Acad. Sci.* **105**, 11766–11771 (2008).

27. J. E. Darnell, Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749 (2002).

28. J. H. Bushweller, Targeting transcription factors in cancer — from undruggable to reality. *Nat. Rev. Cancer* **19**, 611–624 (2019).

29. A. K. Mapp, R. Pricer, S. Sturlis, Targeting transcription is no longer a quixotic quest. *Nat. Chem. Biol.* **11**, 891–894 (2015).

30. I. Landrieu, *et al.*, Characterization of ERM transactivation domain binding to the ACID/PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* **43**, 7110–7121 (2015).

31. S. L. Currie, *et al.*, ETV4 and AP1 Transcription Factors Form Multivalent Interactions with three Sites on the MED25 Activator-Interacting Domain. *J. Mol. Biol.* **429**, 2975–2995 (2017).

32. A. R. Henderson, *et al.*, Conservation of coactivator engagement mechanism enables small-molecule allosteric modulators. *Proc. Natl. Acad. Sci.* **115**, 8960–8965 (2018).

33. G. Loving, B. Imperiali, A Versatile Amino Acid Analogue of the Solvatochromic Fluorophore 4- *N,N* -Dimethylamino-1,8-naphthalimide: A Powerful Tool for the Study of Dynamic Protein Interactions. *J. Am. Chem. Soc.* **130**, 13630–13638 (2008).

34. E. Vojnic, *et al.*, Structure and VP16 binding of the Mediator Med25 activator interaction domain. *Nat. Struct. Mol. Biol.* **18**, 404–409 (2011).

35. I. Radhakrishnan, *et al.*, Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell* **91**, 741–752 (1997).

36. W. C. Pomerantz, *et al.*, Profiling the Dynamic Interfaces of Fluorinated Transcription Complexes for Ligand Discovery and Characterization. *ACS Chem. Biol.* **7**, 1345–1350 (2012).

37. R. Giri, A. Morrone, A. Toto, M. Brunori, S. Gianni, Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc. Natl. Acad. Sci.* **110**, 14942–14947 (2013).

38. S. L. Shammas, A. J. Travis, J. Clarke, Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc. Natl. Acad. Sci.* **111**, 12055–12060 (2014).

39. N. Wang, J. M. Lodge, C. A. Fierke, A. K. Mapp, Dissecting allosteric effects of activator-coactivator complexes using a covalent small molecule ligand. *Proc. Natl. Acad. Sci.* **111**, 12061–12066 (2014).

40. L. Dahal, T. O. C. Kwan, S. L. Shammas, J. Clarke, pKID Binds to KIX via an Unstructured Transition State with Nonnative Interactions. *Biophys. J.* **113**, 2713–2722 (2017).

41. A. Zhuravleva, L. M. Gierasch, Substrate-binding domain conformational dynamics mediate Hsp70 allostery. *Proc. Natl. Acad. Sci.* **112**, E2865–E2873 (2015).

42. S. J. Demarest, *et al.*, Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).

43. C. W. Lee, M. A. Martinez-Yamout, H. J. Dyson, P. E. Wright, Structure of the p53 Transactivation Domain in Complex with the Nuclear Receptor Coactivator Binding Domain of CREB Binding Protein. *Biochemistry* **49**, 9964–9971 (2010).

44. P. Haberz, M. Arai, M. A. Martinez-Yamout, H. J. Dyson, P. E. Wright, Mapping the interactions of adenoviral E1A proteins with the p160 nuclear receptor coactivator binding domain of CBP: Interaction of E1A with the NCB Domain of CBP. *Protein Sci.* **25**, 2256–2267 (2016).

45. G. M. Clore, C. Tang, J. Iwahara, Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr. Opin. Struct. Biol.* **17**, 603–616 (2007).

46. N. J. Anthis, G. M. Clore, Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48**, 35–116 (2015).

47. G. Schreiber, A. E. Keating, Protein binding specificity versus promiscuity. *Curr. Opin. Struct. Biol.* **21**, 50–61 (2011).

48. N. Wang, *et al.*, Ordering a Dynamic Protein Via a Small-Molecule Stabilizer. *J. Am. Chem. Soc.* **135**, 3363–3366 (2013).

49. Z. Wu, *et al.*, Targeting the Transcriptional Machinery with Unique Artificial Transcriptional Activators. *J. Am. Chem. Soc.* **125**, 12390–12391 (2003).

50. L. C. Cesa, *et al.*, X-linked inhibitor of apoptosis protein (XIAP) is a client of heat shock protein 70 (Hsp70) and a biomarker of its inhibition. *J. Biol. Chem.* **293**, 2370–2380 (2018).

51. V. Paakinaho, *et al.*, Single-molecule analysis of steroid receptor and cofactor action in living cells. *Nat. Commun.* **8**, 15896 (2017).

52. F. Delaglio, *et al.*, NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6** (1995).

53. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

54. M. Mobli, A. S. Stern, W. Bermel, G. F. King, J. C. Hoch, A non-uniformly sampled 4D HCC(CO)NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *J. Magn. Reson.* **204**, 160–164 (2010).

Acknowledgements

Financial support for this work was received from NIH R01 GM65530, R35 GM136356(to A.K.M.) and CA207272 (to **UPS Mail Innovations** T.C.). M.J.H. was supported by a fellowship from the Department of Education (Graduate Assistance in Areas of National Need), and B.S.M. by a Michigan Life Sciences Institute Postdoctoral Fellowship. We thank the Mapp lab for helpful comments and D. Sahu for assistance with performing the NMR assignment experiments.

Author Contributions

M.J.H. and A.K.M. conceived and designed research. M.J.H., B.M.L., and B.S.M. performed research. M.J.H. analyzed the data. M.J.H. and A.K.M. wrote the manuscript with input from all coauthors.