



Published in final edited form as:

*Biochem J.* 2011 October 01; 439(1): 15–25. doi:10.1042/BJ20110972.

## Gene-specific transcription activation via long-range allosteric shape-shifting

Chung-Jung Tsai<sup>\*</sup>, Ruth Nussinov<sup>\*,†,1</sup>

<sup>\*</sup>Basic Science Program, SAIC-Frederick, Center for Cancer Research Nanobiology Program NCI-Frederick, Frederick, MD 21702, U.S.A.,

<sup>†</sup>Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

### Abstract

How is specificity transmitted over long distances at the molecular level? REs (regulatory elements) are often far from transcription start sites. In the present review we discuss possible mechanisms to explain how information from specific REs is conveyed to the basal transcription machinery through TFs (transcription factors) and the Mediator complex. We hypothesize that this occurs through allosteric pathways: binding of a TF to a RE results in changes in the AD (activation domain) of the TF, which binds to Mediator and alters the distribution of the Mediator conformations, thereby affecting transcription initiation/activation. We argue that Mediator is formed by highly disordered proteins with large densely packed interfaces that make efficient long-range signal propagation possible. We suggest two possible general mechanisms for Mediator action: one in which Mediator influences PIC (pre-initiation complex) assembly and transcription initiation, and another in which Mediator exerts its effect on the already assembled but stalled transcription complex. We summarize (i) relevant information from the literature about Mediator composition, organization and structure; (ii) Mediator interaction partners and their effect on Mediator conformation, function and correlation to the RNA Pol II (polymerase II) CTD (C-terminal domain) phosphorylation; and (iii) propose that different allosteric signal propagation pathways in Mediator relate to PIC assembly and polymerase activation of the stalled transcription complex. The emerging picture provides for the first time a mechanistic view of allosteric signalling from the RE sequence to transcription activation, and an insight into how gene specificity and signal transmission can take place in transcription initiation.

### Keywords

allosteric pathway; allosteric propagation; allostery; enhancer; mediator; transcription initiation; transcription regulation; transcription start site

---

<sup>1</sup>Corresponding author: ruthnu@helix.nih.gov.

## INTRODUCTION

Cell development, function and survival under stress depend on gene-specific activation and repression. A fundamental question in the life sciences is how does nature accomplish this task? A crucial step is binding of TFs (transcription factors) to DNA REs (response elements). To activate specific genes, TFs need to bind with specific REs and communicate, often over long distances, with RNA Pol II (polymerase II) and the GTFs (general transcription factors, for example TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, TFIIH). However, REs are redundant: for example, a 6 bp RE occurs >700000 times in the human genome, and if we consider RE degeneracy this is probably an underestimation [1]. Thus a TF has to choose among a large number of similar REs across the genome, each associated with a functionally distinct gene. Although many REs are unavailable for TF binding because they are packaged in heterochromatin [2,3], the number of exposed REs is still extremely high. In addition, even though the REs can be near promoters [4], mostly they are at great distances thousands of base pairs away [5]. Thus communication with the transcription machinery needs to be long-range and specific; it should either recruit or move an already assembled but stalled transcription complex off specific promoters. Taken together, this raises two questions: first, how do TFs select specific REs among all of those available in the genome? [1,6]; and secondly, how is the specificity which is encoded in TF-RE binding communicated to the Pol II-GTF machinery to activate gene-specific transcription far away? Insight into selective RE recognition and effective signal transmission, from the binding site of the TF to the RE to a distant Pol II complex, is essential for the understanding of gene-specific activation in different cell types and under fluctuating cellular conditions. We have recently addressed the first question of how TFs select specific REs among all degenerate sequences [1,6]. From the mechanistic standpoint, TF-RE recognition can be understood based on four factors [6-15]: (i) RE availability (i.e. exposed or packed in chromatin); (ii) the physiological state of the cell which fluctuates with the changing environment (protein concentrations, post-translational modification states, the presence of small molecule morphogens etc.); (iii) the dynamic free-energy landscape of the proteins and DNA (i.e. the dynamic distributions of the conformational ensembles) [7,14,15]; and (iv) the tight packing of multiple TFs and co-regulators on regulatory DNA stretches, such as enhancers (which leads to selective RE binding via combinatorial assembly of TFs) [1,6].

In the present review, we address primarily the second question: how is the TF-RE specificity communicated to recruit the PIC (pre-initiation complex), initiate transcription and the subsequent elongation; or, to activate transcription elongation of initiated-and-stalled Pol II-GTFs far away? Communication could take place through interaction between the TF AD (activation domain) and the Pol II-GTF machinery [TFIID and SAGA (Spt-Ada-Gcn5-acetyltransferase) complexes] [16-18], through Mediator which acts as a GTF, at a basal transcription level [16,19] or via an AD-bound 'activated' Mediator conformation. In the present review we focus on the last mechanism.

A transcription event consists of distinct stages, and these are defined and summarized in Figure 1. Within this broad picture, we review recent compelling experimental observations and combine them with current concepts in structural biology to account for two scenarios in the regulation of gene expression at the transcription stage. We organize the data into two

possible mechanisms (Figure 2). In mechanism I, the activated TF-bound Mediator conformation helps in recruiting the Pol II–GTFs, leading to the formation of the PIC (Figure 2A) and transcription initiation and elongation without pausing. In mechanism II, the Pol II–GTFs–Mediator complex [the so-called PEC (pre-elongation complex)] is already promoter-bound in a transcription initiated-and-stalled state (Figure 2B). The binding of the TF AD to PEC triggers Pol II dissociation from the promoter region and the generation of a stable transcript. We propose that different allosteric propagation pathways play key roles in recruiting the PIC and transcription without pausing in mechanism I and in activating the stalled Pol II in mechanism II. Because the events and the location of the AD-binding sites on Mediator differ, the major allosteric propagation pathways in Mediator that are elicited by the AD binding in mechanism I are likely to differ from those in mechanism II. In addition, we propose that the different pathways may also correlate with Ser<sup>5</sup> phosphorylation levels of the Pol II CTD (C-terminal domain).

Although in the present review for both mechanisms we assume that the promoter is exposed and available for binding (Figure 2), allosteric pathways are also likely to be important in the step-by-step assembly of the PEC in mechanism I when the promoter is unavailable, i.e. by TF first binding to the RE, followed by AD recruitment of the chromatin-remodelling complex HAT (histone acetyltransferase), Mediator and GTFs and finally RNA Pol II. Binding events always perturb the structure and the perturbation always propagates through multiple pathways [20], leading to conformational and (or) dynamic changes at other binding sites [21–23]. In the case of the Mediator, if the allosteric conformational change which is elicited by the AD binding is large and takes place in a sufficiently high population of the Mediator conformational ensemble, as could be the case in mechanism II, EM (electron microscopy) images will show ‘shape-shifting’. Unfortunately, currently high-resolution structural data are available only for a small fraction of the Mediator subunits, which precludes mapping of the allosteric pathways. Nonetheless, even though much is still unknown, a coherent framework can help in interpreting observations and guiding experiments assigning functions to Mediator subunits.

## MEDIATOR

Recently, striking clues as to how Mediator can help in signal transmission have been published (e.g. [24–28]) extending earlier observations [17,19,29–32]. Mediator subunits are divided by the modules they fall into: Head, Middle and Tail. There are significant functional differences between yeast and human Mediator; the sequences are poorly conserved and Pol II pausing does not appear to be a significant regulatory step in yeast, in contrast with the more highly developed metazoans where pausing allows a higher degree of regulation and functional complexity. However, from the conformational standpoint, the lack of structural data does not allow us to make this distinction. Table 1 lists some of the known interactions of Mediator’s subunits with TFs. Mediator faces a challenging problem, it has to accurately and efficiently transmit stress, cell type and developmental stage gene-specific RE–TF interactions and post-translational modification events to the Pol II–GTF-binding regions. How does Mediator accomplish this task?

The Head has a jaw [MED (Mediator subunits) 18–20] which moves with respect to subunits MED17 and MED11–22. There is a hinge between MED6 and MED8, which contacts MED18 and MED17 [25] (Figure 3 provides a two-dimensional map of the Mediator subunits, where structural data are available the map follows the yeast data). The open conformation of the Head is stabilized by TBP (TATA-binding protein), and the Pol II subunits Rpb4/Rpb7 bind between the jaws [25]. In addition to the CTD and Rpb4/Rpb7, the Rpb3 subunit of Pol II has direct interaction points with Mediator at MED17 [33]. Comparative genomics [34] and sequence analysis of Mediator subunits indicated that the Tail is the least conserved module [35], suggesting a more recent emergence. The Head is the most conserved module suggesting its greater age. It is worth noting that based on bioinformatics analysis it has been found that: (i) the conformational disorder level of the Mediator exceeds that in other complexes of similar size [35]; (ii) disorder is particularly high in the Tail and Middle modules; (iii) the extent of disorder increases from yeast to human; and (iv) the disordered regions arrangements and interaction sites are similar between yeast and human implicating specific recognition at protein–protein interfaces [35], which is corroborated by crystal structure data for the MED8–MED18–MED20 [36–37], MED7–MED31, MED7–MED21 [38] and MED11–MED22 [39] subunits (Figure 4). Taken together with EM images, these structures suggest that the surfaces available for TF binding are largely in the Mediator Tail, Head and in the Middle subunit MED1 (Figure 3). This is in agreement with biochemical interaction data [24]. Otherwise, much of the Middle subunits surface is occupied by other subunits. The structures also emphasize the shortcomings of two-dimensional Mediator maps: although these suggest apparent limited MED7–MED21 interactions [35,40,41], the MED7–MED21 interface spans one third of the Mediator length [38] with ~150 conformationally disordered (thus not crystallized) MED7 residues which provide an additional large interaction area.

A key mechanistic question is then how TF–Mediator binding activates Pol II and how is this related to the requirement of CTD phosphorylation?

## **MEDIATOR UNDERGOES ‘SHAPE-SHIFTING’: A VIEW FROM THE STANDPOINT OF THE FREE ENERGY LANDSCAPE**

Recently, Taatjes and colleagues [27] presented compelling data illustrating how p53 can activate transcription via Mediator ‘structural shifts’, extending earlier key observations which indicated a dramatic conformational change upon TF binding [28–30]. The authors showed that the p53AD and the p53CTD interact with MED17 and MED1 respectively. The interactions differentially affect Mediator structure and Pol II activity, whereas the p53AD induces a conformational change and this correlates with activation of stalled Pol II, this is not the outcome of the p53CTD–MED1 interaction. Furthermore, mutations of p53AD residues (L22Q and W23S), which prevent expression of most p53 target genes [42,43], disrupted the MED17 interaction. They concluded that Mediator undergoes p53AD-induced shape-shifting, which is expressed in Mediator–Pol II binding. This work is important because it allows a first glimpse of the entire communication pathway: how TF bound to a RE far away from the promoter can activate Pol II. Shape-shifting is a fundamental consequence of binding events, which allosterically redistribute conformational ensembles

[14,15,44–47]. The single most important principle, based on statistical mechanics, observed by NMR and described in terms of the free energy landscape, is that all (dynamic) proteins exist in conformational ensembles around their native states [14,15,44–49]. The states present slightly different conformations separated by low barriers. Perturbations such as those caused by cellular binding events (by proteins, DNA, small molecules etc.), covalent changes (for example post-translational modifications such as phosphorylation, acetylation, methylation, ubiquitination and mutational events), protein concentrations, temperature changes, ionic strength and pH, alter the relative distributions of the states. This change in the distribution is called ‘population shift’ [14,15,44–47]. The strain energy created by the perturbation is released when the atoms rearrange. Population shift takes place because the strain energy spreads, that is, the rearrangement propagates in the molecular structure through the dynamic changes in atomic contacts. Allostery is such an event and the agents listed above are allosteric effectors. To understand allosteric propagation, it is important to consider the entire ensemble because it is the population that shifts. An ensemble consists of states which are characterized by unique combinations of some structural parameters, typically backbone dihedral angles and side-chain rotamers each with an assigned energy level. These characteristics can define the degrees of freedom (possible ‘microstates’) of the protein. The conformational entropy associated with a particular conformation in the ensemble depends on the probability that the state will be occupied. The higher the temperature, the larger is the number of accessible states (called ‘thermally accessible states’). States which are occupied at higher temperatures are rare under physiological conditions, but their populations typically increase during allosteric propagation. The changes in the conformations can be very minor [22,46,50] and expressed only in subtle changes in side chains and backbones, or in conformational dynamics [22], and as such not observed by comparisons of crystal structures as in the case for the p53 DNA-binding domain [51–52]. Allosteric propagation can be described as a progressive change in the occupancy of states. In this way proteins also use rare states populated only at high temperature to execute their function, which play key roles in molecular recognition (reviewed in [47]) and in transmission of information [12]. As such they are also likely to be conserved by evolution.

In our case, in the present review, DNA is an allosteric effector: slightly different RE sequences can have slightly different contacts with the TF surface atoms and these changes propagate and can lead to conformations of TF that present altered surfaces of binding sites far away. This is in particular the case if those binding sites are in conformationally disordered domains which even on their own present conformational heterogeneity with low barriers separating the states, as in the p53AD [53].

In turn, the altered surfaces may recognize and interact with different Mediator subunits. Because new interactions are formed which involve residues on the surface of Mediator, changes in Mediator atomic contacts will take place, either between Mediator atoms or between Mediator and the solvent. These changes perturb the structure of the Mediator-binding site. The changes proceed through different pathways (thus subunits) throughout the Mediator structure, and will reach the Pol II/GTF-binding sites. The conformational states with the altered binding sites now become more populated. Currently, allostery is overlooked frequently when trying to understand how binding phenomena determine the functional

outcome. For the glucocorticoid receptor, a change of even a single base pair between REs led to altered TF conformation away from the binding site which expresses a different function [10,11].

## ALLOSTERIC SIGNAL PROPAGATION PATHWAYS IN MEDIATOR AND POL II ACTIVATION

Because some gene activation events require high Pol II CTD (Ser<sup>5</sup>) phosphorylation levels whereas others do not, the role of phosphorylation in the initiation of transcription has been unclear [27]. We argue that consideration of allosteric pathways from the TF-binding site to Pol II can help, and suggest that the major allosteric propagation pathways from Mediator cross into Pol II at two locations which we call ‘triggering points’: the Rpb4/Rpb7-binding region in the Mediator Head and around the CTD-interaction site in Middle subunit MED26 (Figure 3). We consider the allosteric effects in two possible mechanisms. In mechanism I, the TF-bound Mediator recruits Pol II–GTF leading to formation of the PIC and transcription initiation and elongation without stalling (depicted in Figure 1A). In mechanism II, PEC is already promoter-bound in a transcription initiated-and-stalled state. In mechanism I, TF can bind to Tail/Middle subunits. TF binding can help in opening the Pol II-binding pocket in Mediator by allosterically increasing the open-state population, which will facilitate the recruitment of Pol II. In addition, a direct interaction of MED11 and the Rad3 subunit of TFIIF facilitate phosphorylation of CTD at Ser<sup>5</sup> by CDK (cyclin-dependent kinase) 7 of TFIIF [54]. The resulting phosphorylation pattern helps to recruit the RNA-capping enzyme, which adds a 5' cap to nascent RNA transcripts [55,56]. We suggest that the phosphorylation of the Pol II CTD at Ser<sup>5</sup> allosterically prevents the binding of the NELF (negative elongation factor; see below), which leads to transcription activation without pause. Thus TF binding and activation of Mediator initiates the formation of the PIC. On the other hand, in mechanism II, the activated Mediator does not help to form a PIC. Therefore there is low CTD phosphorylation at Ser<sup>5</sup>. Under such circumstances, the recruitment of the RNA-CE (capping enzyme) to the stalled PEC, which is needed to fulfil the initiation of elongation checkpoint, can be via the unphosphorylated CTD of Spt5 [56]. The association of NELF with Spt5 and Pol II enables the binding of the NELF-E subunit to the newly synthesized RNA. This prevents any further extrusion of the nascent RNA which stalls the PEC [57]. Although there are several possible explanations for PEC stalling at the proximal promoter region [57], we attribute gene regulation in mechanism II to the NELF-mediated pausing.

The dissociation of Pol II requires energy. If the TF-binding site on Mediator is far from the interaction site with Pol II, the strain energy generated by the TF binding will dissipate in the gigantic Mediator complex and the amount reaching the Pol II-binding site will be insufficient. In mechanism II, sufficient energy may reach a nearby Pol II ‘triggering point’ to overcome barriers even at low phosphorylation states. The p53AD interaction with the Head MED17 [27] and phosphorylated ELK1 (Ets-like transcription factor 1)-binding Tail MED23 [24] provide examples. ELK1 activation of initiated-and-paused Pol II at the Egr1 promoter [24] may constitute a particularly nice example of the relevance of allosteric pathways. Even for the same gene different REs can lead to altered pathways in Mediator, as



suggested by the MED23 requirement when ELK1 binds one RE in embryonic stem cells, but not when binding another in fibroblasts. On the other hand, the energy that reaches a triggering point is insufficient for Pol II dissociation in mechanism I. VP16 which interacts with Tail MED25 [58–59], and PPAR  $\gamma$  (peroxisome-proliferator-activated receptor  $\gamma$ ) with MED14 [26] require higher phosphorylation levels and provide examples. Mutational studies (e.g. p53AD [27] or a subunit knock-out [24] which could abolish propagation) would assist in mapping major propagation pathways [20]. The observation that the CDK8 module may function as a stimulus-specific transcription activator for p53 at the p21 locus (mechanism II) in the presence of Nutlin, but not in response to cell stress upon UV irradiation (mechanism I), could relate to propagation [60]. Cyclin C, MED12 and CDK8 may not be required for p21 where the p53AD binds MED17 in the Head module, and can be recruited only following Nutlin treatment. This may not to be the case under UV light C. Previous studies suggested that mechanism II is less frequent. Chromatin immunoprecipitation assays coupled with genomic microarray experiments in *Drosophila* and mammalian cells observed only 20–30% enriched Pol II populations at the 5' ends of genes [61–65]; however, more recent results have suggested a much higher occurrence [66,67].

Above, we have highlighted the role of allosteric propagation through Mediator in transcription initiation/elongation. Yet, Mediator is a huge complex. The question arises of how do the signals transmit effectively over such distances?

## THE ROLE OF CONFORMATIONAL DISORDER IN MEDIATOR SUBUNITS IN EFFECTIVE SIGNAL PROPAGATION

Disordered proteins lack a strong hydrophobic core and consequently they exist in solution in two states, unbound and disordered or bound and folded with the hydrophobic core forming at their interfaces [68]. Upon binding, they undergo disorder-to-order transition [68,69]. They bind with large tight functionally required interfaces [70–73]. Similar to the hydrophobic cores of stable single-chain proteins, such well-packed interfaces imply high affinity and specificity. The fact that EM images show particles validates high subunit organizational specificity (Figure 4). Because signals are transmitted in the assembly through breaking and forming atom–atom contacts within and across subunits, the large densely packed interfaces facilitate efficient propagation. Essentially, the multi-subunit Mediator resembles a single protein. Subunit perturbation by, for example, TF binding, acetylation or phosphorylation (as in the MED1 subunit [74]) can allosterically propagate efficiently. The huge Mediator size {when unbound to Pol II the estimated height is  $\sim 400$  Å ( $1$  Å =  $0.1$  nm) with a triangular shape [75] and upon binding the dramatic conformational change to elongated shape suggests longer distances requires effective signalling between binding sites which are far away. Such a view of disordered states can hold for signalling proteins [76,78], which are also overwhelmingly disordered. The Mediator and GTFs sizes (together over  $\sim 70$  subunits in the human [79]) can explain why disorder is a key property: disorder implies highly specific protein–protein interactions and efficient signalling over large distances.

## ASSEMBLED POL II–GTFs CAN ALSO PRESENT DIFFERENT INTERACTION SURFACES

In addition to Mediator, Pol II–GTFs recruited or stalled in various developmental gene- and tissue-specific environments can also present different interaction surfaces. Previous analyses have shown that: (i) the TATA box is not always present and TBPs are not a general requirement for gene activation; (ii) TBP-related factors can mediate Pol II transcription initiation; and (iii) TBP or TBP-related factors are not necessarily a prerequisite for transcriptional activity. Instead, bioinformatics and experimental results point to combinatorial sets of core promoters [80] and references therein). This suggests that pre-initiation or initiated-and-paused complexes can have diverse effects on GTF composition depending on promoter sequence elements and their spacing [1]. For example, the TFIID (14–16 subunits) [79,80] may lack some TAFs (TBP-associated factors) and associate with other TAF combinations. Not all TAFs are universally required. Some (like TAF10) may differentially populate core promoters in cell types and developmental stages [80]. Genome-wide analysis indicated a good correlation (75%) between TAF1 occupancy and gene activity [63]. The remaining 25% were either active without TAF1 binding, or TAF1-bound but with no detected transcription.

## MEDIATOR-ASSISTED AND UNASSISTED SCENARIOS

In the scenarios described above, we have assumed that the promoters are available. Under such circumstances, if the concentration of Pol II and GTFs is high, all promoters will be occupied. If they are relatively low and the Pol II–GTF affinity to the promoter is not high, the Mediator may help in recruitment. In a second scenario, the promoters are unavailable; in such a case, cofactors such as HATs (for example PCAF {p300/CBP [CREB (cAMP-response-element-binding protein)-binding protein]}) may help in chromatin remodelling. In a third scenario (which is not discussed in the present review), there is no Mediator involvement. RE-bound TFs and their cofactors may recruit Pol II–GTFs [81,82] or function in exposing the promoter. Enhanceosomes [4] consist of clusters of REs, TFs and their cofactors, some of which have enzymatic activity similar to the Mediator subunits. From the conformational standpoint, the mechanism is similar: signal propagation elicited by binding or post-translational modification events. Given the (frequent) large distances between enhancers and promoters, Mediator-assisted scenarios appear to be the major mechanism. Under such Mediator-assisted scenarios, mechanism I will take place if: (i) the promoter is unavailable; (ii) the Pol II/GTF concentration is low; or (iii) promoter recognition sequences present low affinity to Pol II/GTFs. The prevalence of mechanism I (~70–80% [61–65]) suggests that these are common occurrences, and are probably due to the first and third factors as illustrated by the failure of bioinformatics approaches to detect consensus sequences.



## HOW IS SPECIFICITY COMMUNICATED FROM THE RE TO THE POL II –GTF MACHINERY?

To summarize, mechanistically, the pathway starting from the RE and reaching the Pol II–GTF can be divided into the steps below. (i) TF binds available REs [1,6,8,9,83]. If the TF concentration is high, it binds all REs; if low, only high-affinity REs. Similar REs have slightly different atomic interactions with the TF, which elicit altered distributions of allosteric pathways [20] and thus different TF surfaces. For the p53 example, although REs interact with the DNA-binding domain, the conformational changes can be expressed in the p53-binding site to ASPPs (apoptosis-stimulating proteins) [84], the p53AD and p53CTD [27]. (ii) The conformationally altered TF AD surfaces recruit either different Mediator subunits, or conformational states of the same subunit (phosphorylated, acetylated etc.). Initially, the ensemble of Mediator conformations is largely in the closed state. TF binding to the Tail/Middle modules induces allosteric effects in Mediator subunits (for example MED14 and MED1; Figure 4). Pathways propagate through other Mediator subunits (e.g. MED4, MED9, MED7 and MED21) increasing the populations of Mediator open states [15,6], which enhances Pol II recruitment. Binding different subunits or conformational states of the same Mediator subunit can elicit altered distribution of allosteric pathways [20]. Even if other TFs bind the same subunits (for example, thyroid hormone receptor  $\alpha$ , glucocorticoid receptor, PPAR  $\gamma$  and oestrogen receptor  $\alpha$ , all bind MED1) or the same TFs bind different Mediator subunits (for example glucocorticoid receptor binds with MED1 [85–86], MED14 [85] and MED15 [87]; VP16 binds with MED17 [88], MED15 [89] and MED25 [89]), the interactions will elicit different allosteric pathways. In the absence of binding of the TF AD, Mediator can still recruit the Pol II machinery at a basal rate, because a low population of the open states exists in solution. These will bind via a conformational selection followed by population shift mechanism [14,45,90].

Hence, in effect, there is specific long-range allosterically induced signal propagation from the RE (an allosteric effector) through the TF, Mediator subunits and finally Pol II. Because allostery proceeds through multiple (major and minor) pathways, propagation at each step depends on the perturbation at previous steps.

## A BROADER OUTLOOK OF ALLOSTERY IN TRANSCRIPTION INITIATION AND ELONGATION AND OPEN QUESTIONS

Recent studies by Cramer and colleagues [91] and Kornberg and colleagues [92] have provided structural details of the Pol II–TFIIB complex in different transcription initiation states and offer an unprecedented insight that illuminates the role of TFIIB in the closed-to-open promoter transition. The GTFs locate the promoter and (with possible help from Mediator) open the DNA (TFIIH is a helicase). TFIIB recruits Pol II, and the TFIIB tunnel stabilizes the melted region ~20 bp downstream of the TATA box by interacting with the template DNA strand. Once a ~7-residue-long transcript is synthesized [92], an 8 bp DNA stretch at the bubble head re-anneals. The TFIIB B finger reaches into the Pol II active centre and clashes with a longer RNA transcript and this leads to promoter escape. In mechanism I and in basal transcription [19] there is no pausing in transcript elongation; in

mechanism II there is. The question arises as to what leads to the transition into processive elongation of the initiated-and-paused complex in mechanism II and why no pausing in mechanism I.

In both mechanisms, the addition of the 7-methyl-guanosine 'cap' to the 5' end of the nascent RNA transcript by a CE and an RNMT (RNA methyltransferase) [93] is a critical checkpoint for RNA elongation. The subsequent transition from elongation initiation into processive elongation (Figure 1A) requires the recruitment of the p-TEFb (positive transcription elongation factor b) which contains CycT1 (cyclin T1) or CycT2 (cyclin T2) and CDK9. We suggest that CE and RNMT recruitment differs in the two mechanisms. In mechanism I (Figure 2A), PEC assembly is initiated by the TF AD binding to Mediator. Mediator allosterically stimulates CDK7 (a subunit of TFIIH), which phosphorylates Pol II CTD (at Ser<sup>5</sup> in the heptapeptide repeats). The phosphorylated CTD recruits CE and RNMT which help recruiting p-TEFb without interruption. Therefore there is no pausing in mechanism I. In mechanism II (Figure 2B), PEC assembles at the promoter without involvement of the TF AD, thus with a low level of Pol II CTD phosphorylation [27]. Low CTD phosphorylation can lead to CE recruitment by the unphosphorylated C-terminal repeat of transcription elongation factor Spt5 which binds to the Pol II Clamp [94–97] and recruits NELF. NELF associates with the nascent capped RNA transcript and prevents RNMT from recruiting p-TEFb if RNMT was recruited by Spt5, but not if it was recruited by Pol II. Therefore NELF binding in mechanism II will stall the elongation of RNA transcripts (unless some alternative route is available for p-TEFb recruitment [57,98]). However, without the impediment of NELF bound to Pol II, there is no substantial delay in transcription from elongation initiation to processive elongation. We propose that binding of the AD to Mediator Head (as in the p53AD) leads to a significant conformational change which may allosterically propagate via MED17/MED18/MED20 to Rpb4/Rpb7 and the Pol II Clamp and may dislocate NELF. This allows p-TEFb recruitment via RNMT to drive the stalled PEC into a processive transcript elongation state [57,99].

Allosteric propagation (Figure 3): (i) offers an explanation why MED14 and MED7/MED21 are essential; in contrast, the essential role of MED17 [19] is not related to allostery; (ii) further suggests that Pol II CTD phosphorylation is unlikely to take place in basal transcription (where there is no TF involvement), or if the TFAD binds to the Mediator Head (thus the conformational change is sufficiently large); and (iii) also suggests that despite the increasing number of observed unphosphorylated CTD cases in initiated-and-stalled transcription activation, in the majority of Mediator-assisted cases, Pol II CTD phosphorylation will take place. Mutational and gene knockout experiments may be expected to identify Mediator subunits whose function is to mediate allosteric pathways.

## CONCLUSIONS

In the present review, we have organized the experimental observations into two possible mechanisms. In mechanism I TF-bound Mediator helps in recruiting Pol II–GTFs (Figure 2A) and in mechanism II Pol II – GTFs–Mediator is in a promoter-bound transcription initiated-and-stalled state prior to TF involvement (Figure 2B). In mechanism I activation of recruited Pol II could be CTD phosphorylation-dependent, which is not the case in

mechanism II. From the evolutionary standpoint, mechanism II appears older as the Head is more conserved with a lower disorder level [35]. Pol II recruitment (mechanism I) requires the Head, Middle and Tail. In mechanism II, the Head alone might be sufficient. We suggest that mechanism I requires additional cofactors and in this way allows finer regulatory control. p53 activation of p21 provides one example: Nutlin (mechanism II) does not require Middle submodule CDK8, unlike UV light C (mechanism I) which does [60]. MED23 knockouts [24] present an example where allosteric signalling from different REs of the same gene propagate via different pathways. Egr1 activation by ELK1 in embryonic stem cells requires MED23, which is not the case in fibroblasts. Initiation of elongation of initiated-and-stalled PEC (mechanism II) is via MED23, a Tail subunit, suggesting a possible interaction between MED23 and a Head subunit. Mechanism II also appears to be non-essential. The Head subunits MED18 and MED20 which contact Pol II Rpb4/Rpb7 are non-essential for viability in yeast [100]. A preassembled and stalled PEC is advantageous, it saves assembly time under cellular stress, such as at higher temperature when accelerated heat-shock protein expression is needed [95]. Similar to mechanism II, there is no Pol II CTD phosphorylation in basal transcription and the required MED17 (for PEC formation [100]) is in Mediator Head [19]; however, the mechanisms are different. In mechanism II, following TF binding, signals propagate to Pol II Rpb4/Rpb7 via MED18/MED20 which show shape-shifting [27], whereas there is no TF activation in basal transcription.

To conclude, recent observations combine to paint a coherent scheme of how specific gene activation from the RE to Pol II activation can take place. They suggest how a RE-bound TF changes its conformation [10,11], and communicates with Pol II–GTFs via Mediator to change their shapes [25,27] and facilitate Pol II promoter clearance. When we consider that changes in molecular shape are the outcome of structural perturbation, that RE specificity is a key factor in specific gene activation [8,9] and that Mediator subunits present an unusual disorder level [35], the overall picture clarifies. Ample data support the paramount importance of specific RE recognition in determining functional consequences [1,6,8–10]. RE–TF specificity can propagate via large compact interfaces between Mediator subunits, culminating in specific Mediator ‘shape-shifts’ and Pol II interactions. Combined with cellular conditions which govern protein concentrations and their post-translational modification states [6], epigenetic events [2,3] and general TFs assembling on core promoters [80,101] may provide a framework to understand how specific far-away REs can specify gene activation. This leads us to describe Mediator functions as including: (i) increasing the effective local concentration of Pol II and GTFs [91]; (ii) regulation of the PEC; (iii) transmitting and funnelling signals from REs and post-translational modification events (reflecting cellular conditions) to the Pol II–GTF machinery binding sites; and (iv) helping to activate a stalled Pol II. Functional assignment of Mediator subunits could uncover Mediator subunits whose role is signal transmission to control genome-wide transcription regulation. This could be the case for MED14 and MED7/MED21 through which propagation pathways, which are elicited by TFs binding to Tail or Middle subunits (mechanism I), could go (Figure 3). Function is performed by molecular assemblies that change their shape and dynamics in the presence of an input molecule and this shape change can turn on or off catalytic activity.

Shape-shifting has already been related to allosteric effects [102]. A change of shape is brought about by energy transfer from a perturbation site through atomic contacts, dissipating in the molecule-like waves. If the change of shape is sufficiently large and is displayed by a sufficiently high population, as in Mediator following p53AD interaction with the MED17 subunit in the Head module near the Pol II Rpb4/Rpb7-binding site, it will be observed in EM images. Allosteric signal propagation is the only way that information can propagate. The question is whether the energy reaching the target site is large enough to lead to the required conformational change or an added source, such as phosphorylation which would be recognized by another protein, is needed.

## ACKNOWLEDGEMENTS

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

### FUNDING

The work of the authors has been funded in whole or in part by the National Cancer Institute under contract number HHSN261200800001E. This research was supported (in part) by the Intramural Research Program of the National Cancer Institute Center for Cancer Research.

## Abbreviations used:

<b>AD</b>	activation domain
<b>CDK</b>	cyclin-dependent kinase
<b>CE</b>	capping enzyme
<b>CTD</b>	C-terminal domain
<b>CycT1</b>	cyclin T1
<b>CycT2</b>	cyclin T2
<b>ELK1</b>	Ets-like kinase 1
<b>EM</b>	electron microscopy
<b>GTF</b>	general transcription factor
<b>HAT</b>	histone acetyltransferase
<b>MED</b>	Mediator subunit
<b>NELF</b>	negative elongation factor
<b>PEC</b>	pre-elongation complex
<b>PIC</b>	pre-initiation complex
<b>Pol II</b>	polymerase II
<b>RE</b>	regulatory element

<b>PPAR<math>\gamma</math></b>	peroxisomeproliferator-activated receptor $\gamma$
<b>p-TEFb</b>	positive transcription elongation factor b
<b>RNMT</b>	RNA methyltransferase
<b>TBP</b>	TATA-binding protein
<b>TAF</b>	TBP-associated factor
<b>TF</b>	transcription factor

## REFERENCES

- Pan YP, Tsai CJ, Ma BY and Nussinov R (2010) Mechanisms of transcription factor selectivity. *Trends Genet.* 26, 75–83 [PubMed: 20074831]
- Segal E and Widom J (2009) What controls nucleosome positions? *Trends Genet.* 25, 335–343 [PubMed: 19596482]
- Hager GL, McNally JG and Misteli T (2009) Transcription Dynamics. *Mol. Cell* 35, 741–753 [PubMed: 19782025]
- Panne D (2008) The enhanceosome. *Curr. Opin. Struct. Biol* 18, 236–242 [PubMed: 18206362]
- Hakim O, Sung MH and Hager GL (2010) 3D shortcuts to gene regulation. *Curr. Opin. Cell Biol* 22, 305–313 [PubMed: 20466532]
- Pan YP, Tsai CJ, Ma BY and Nussinov R (2009) How do transcription factors select specific binding sites in the genome? *Nat. Struct. Mol. Biol* 16, 1118–1120 [PubMed: 19888307]
- Ma BY, Tsai CJ, Pan YP and Nussinov R (2010) Why does binding of proteins to DNA or proteins to proteins not necessarily spell function? *ACS Chem. Biol* 5, 265–272 [PubMed: 20151694]
- Riley T, Sontag E, Chen P and Levine A (2008) Transcriptional control of human p53-regulated genes. *Nat. Rev. Mol. Cell Biol* 9, 402–412 [PubMed: 18431400]
- Menendez D, Inga A and Resnick MA (2009) The expanding universe of p53 targets. *Nat. Rev. Cancer* 9, 724–737 [PubMed: 19776742]
- Meijsing SH, Pufall MA, So AY, Bates DL, Chen L and Yamamoto KR (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* 324, 407–410 [PubMed: 19372434]
- Gronemeyer H and Bourguet W (2009) Allosteric effects govern nuclear receptor action: DNA appears as a player. *Sci. Signal* 2, pe34 [PubMed: 19491383]
- Gardino AK, Villali J, Kivenson A, Lei M, Liu CF, Steindel P, Eisenmesser EZ, Labeikovsky W, Wolf-Watz M, Clarkson MW and Kern D (2009) Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell* 139, 1109–1118 [PubMed: 20005804]
- Fraser JS, Clarkson MW, Degnan SC, Erion R, Kern D and Alber T (2009) Hidden alternative structures of proline isomerase essential for catalysis. *Nature* 462, 669–673 [PubMed: 19956261]
- Tsai CJ, Kumar S, Ma BY and Nussinov R (1999) Folding funnels, binding funnels, and protein function. *Protein Sci.* 8, 1181–1190 [PubMed: 10386868]
- Gunasekaran K, Ma BY and Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct. Funct. Genet. Bioinfo* 57, 433–443
- Ansari SA, He QY and Morse RH (2009) Mediator complex association with constitutively transcribed genes in yeast. *Proc. Natl. Acad. Sci. U.S.A* 106, 16734–16739 [PubMed: 19805365]
- Kornberg RD (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci* 30, 235–239 [PubMed: 15896740]
- Taatjes DJ, Marr MT and Tjian R (2004) Regulatory diversity among metazoan co-activator complexes. *Nat. Rev. Mol. Cell Biol* 5, 403–410 [PubMed: 15122353]
- Takagi Y and Kornberg RD (2006) Mediator as a general transcription factor. *J. Biol. Chem* 281, 80–89 [PubMed: 16263706]

20. del Sol A, Tsai CJ, Ma BY and Nussinov R (2009) The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure* 17, 1042–1050 [PubMed: 19679084]
21. Tzeng SR and Kalodimos CG (2009) Dynamic activation of an allosteric regulatory protein. *Nature* 462, 368–372 [PubMed: 19924217]
22. Tsai CJ, del Sol A and Nussinov R (2008) Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* 378, 1–11 [PubMed: 18353365]
23. Popovych N, Sun SJ, Ebright RH and Kalodimos CG (2006) Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* 13, 831–838 [PubMed: 16906160]
24. Balamotis MA, Pennella MA, Stevens JL, Wasylyk B, Belmont AS and Berk AJ (2009) Complexity in transcription control at the activation domain-mediator interface. *Sci. Signal* 2, ra20 [PubMed: 19417216]
25. Cai G, Imasaki T, Yamada K, Cardelli F, Takagi Y and Asturias FJ (2010) Mediator Head module structure and functional interactions. *Nat. Struct. Mol. Biol.* 17, 273–279 [PubMed: 20154708]
26. Grontved L, Madsen MS, Boergesen M, Roeder RG and Mandrup S (2010) MED14 tethers mediator to the N-terminal domain of peroxisome proliferator-activated receptor  $\gamma$  and is required for full transcriptional activity and adipogenesis. *Mol. Cell. Biol.* 30, 2155–2169 [PubMed: 20194623]
27. Meyer KD, Lin SC, Bernecky C, Gao YF and Taatjes DJ (2010) p53 activates transcription by directing structural shifts in Mediator. *Nat. Struct. Mol. Biol.* 17, 753–760 [PubMed: 20453859]
28. Taatjes DJ, Naar AM, Ansel F, Nogales E and Tjian R (2002) Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* 295, 1058–1062 [PubMed: 11834832]
29. Asturias FJ (2004) RNA polymerase II structure, and organization of the preinitiation complex. *Curr. Opin. Struct. Biol.* 14, 121–129 [PubMed: 15093825]
30. Davis JA, Takagi Y, Kornberg RD and Asturias FJ (2002) Structure of the yeast RNA polymerase II holoenzyme: mediator conformation and polymerase interaction. *Mol. Cell* 10, 409–415 [PubMed: 12191485]
31. Kang JS, Kim SH, Hwang MS, Han SJ, Lee YC and Kim YJ (2001) The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* 276, 42003–42010 [PubMed: 11555651]
32. Takagi Y, Calero G, Komori H, Brown JA, Ehrensberger AH, Hudmon A, Asturias F and Kornberg RD (2006) Head module control of mediator interactions. *Mol. Cell* 23, 355–364 [PubMed: 16885025]
33. Soutourina J, Wydau S, Ambroise Y, Boschiero C and Werner M (2011) Direct interaction of RNA Polymerase II and Mediator required for transcription in vivo. *Science* 331, 1451–1454 [PubMed: 21415355]
34. Bourbon HM (2008) Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.* 36, 3993–4008 [PubMed: 18515835]
35. Toth-Petroczy A, Oldfield CJ, Simon I, Takagi Y, Dunker AK, Uversky VN and Fuxreiter M (2008) Malleable machines in transcription regulation: the Mediator complex. *PLoS Comp. Biol.* 4, e1000243
36. Shaikhibrahim Z, Rahaman H, Wittung-Stafshede P and Bjorklund S (2009) Med8, Med18, and Med20 subunits of the Mediator head domain are interdependent upon each other for folding and complex formation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20728–20733 [PubMed: 19934057]
37. Larivière L, Geiger S, Hoepfner S, Rother S, Strasser K and Cramer P (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat. Struct. Mol. Biol.* 13, 895–901 [PubMed: 16964259]
38. Baumli S, Hoepfner S and Cramer P (2005) A conserved mediator hinge revealed in the structure of the MED7 center dot MED21 (Med7 Srb7) heterodimer. *J. Biol. Chem.* 280, 18171–18178 [PubMed: 15710619]
39. Seizl M, Larivière L, Pfaffeneder T, Wenzek L and Cramer P (2011) Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. *Nucleic Acids Res.* 39, 6291–6304 [PubMed: 21498544]
40. Lodish HF (2008) *Molecular Cell Biology*, W.H. Freeman, New York



41. Guglielmi B, van Berkum NL, Klapholz B, Bijma T, Boube M, Boschiero C, Bourbon HM, Holstege FCP and Werner M (2004) A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res* 32, 5379–5391 [PubMed: 15477388]
42. Jimenez GS, Nister M, Stommel JM, Beeche M, Barcarse EA, Zhang XQ, O’Gorman S and Wahl GM (2000) A transactivation-deficient mouse model provides insights into Trp53 regulation and function. *Nat. Genet* 26, 37–43 [PubMed: 10973245]
43. Lin JY, Chen JD, Elenbaas B and Levine AJ (1994) Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. *Genes Dev.* 8, 1235–1246 [PubMed: 7926727]
44. Tsai CJ, Ma BY and Nussinov R (1999) Folding and binding cascades: shifts in energy landscapes. *Proc. Natl. Acad. Sci. U.S.A* 96, 9970–9972 [PubMed: 10468538]
45. Kumar S, Ma BY, Tsai CJ, Sinha N and Nussinov R (2000) Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* 9, 10–19 [PubMed: 10739242]
46. Tsai CJ, Del Sol A and Nussinov R (2009) Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Mol. Biosyst* 5, 207–216 [PubMed: 19225609]
47. Boehr DD, Nussinov R and Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol* 5, 789–796 [PubMed: 19841628]
48. Dill KA (1999) Polymer principles and protein folding. *Protein Sci.* 8, 1166–1180 [PubMed: 10386867]
49. Frauenfelder H, Sligar SG and Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254, 1598–1603 [PubMed: 1749933]
50. Daily MD and Gray JJ (2007) Local motions in a benchmark of allosteric proteins. *Proteins Struct. Funct. Genet. Bioinfo* 67, 385–399
51. Kitayner M, Rozenberg H, Rohs R, Suad O, Rabinovich D, Honig B and Shakked Z (2010) Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol* 17, 423–429 [PubMed: 20364130]
52. Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE and Shakked Z (2006) Structural basis of DNA recognition by p53 tetramers. *Mol. Cell* 22, 741–753 [PubMed: 16793544]
53. Wells M, Tidow H, Rutherford TJ, Markwick P, Jensen MR, Mylonas E, Svergun DI, Blackledge M and Fersht AR (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U.S.A* 105, 5762–5767 [PubMed: 18391200]
54. Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, Holstege F and Werner M (2008) Mediator-dependent recruitment of TFIID modules in preinitiation complex. *Mol. Cell* 31, 337–346 [PubMed: 18691966]
55. Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ and Reinberg D (2004) Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A* 101, 7572–7577 [PubMed: 15136722]
56. Schneider S, Pei Y, Shuman S and Schwer B (2010) Separable functions of the fission yeast Spt5 carboxyl-terminal domain (CTD) in capping enzyme binding and transcription elongation overlap with those of the RNA Polymerase II CTD. *Mol. Cell. Biol* 30, 2353–2364 [PubMed: 20231361]
57. Gilmour D (2009) Promoter proximal pausing on genes in metazoans. *Chromosoma* 118, 1–10 [PubMed: 18830703]
58. Mittler G, Stuhler T, Santolin L, Uhlmann T, Kremmer E, Lottspeich F, Berti L and Meisterernst M (2003) A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* 22, 6494–6504 [PubMed: 14657022]
59. Yang FJ, DeBeaumont R, Zhou S and Naar AM (2004) The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci. U.S.A* 101, 2339–2344 [PubMed: 14983011]
60. Donner AJ, Szostek S, Hoover JM and Espinosa JM (2007) CDK8 is a stimulus-specific positive coregulator of p53 target genes. *Mol. Cell* 27, 121–133 [PubMed: 17612495]
61. Core LJ and Lis JT (2008) Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319, 1791–1792 [PubMed: 18369138]

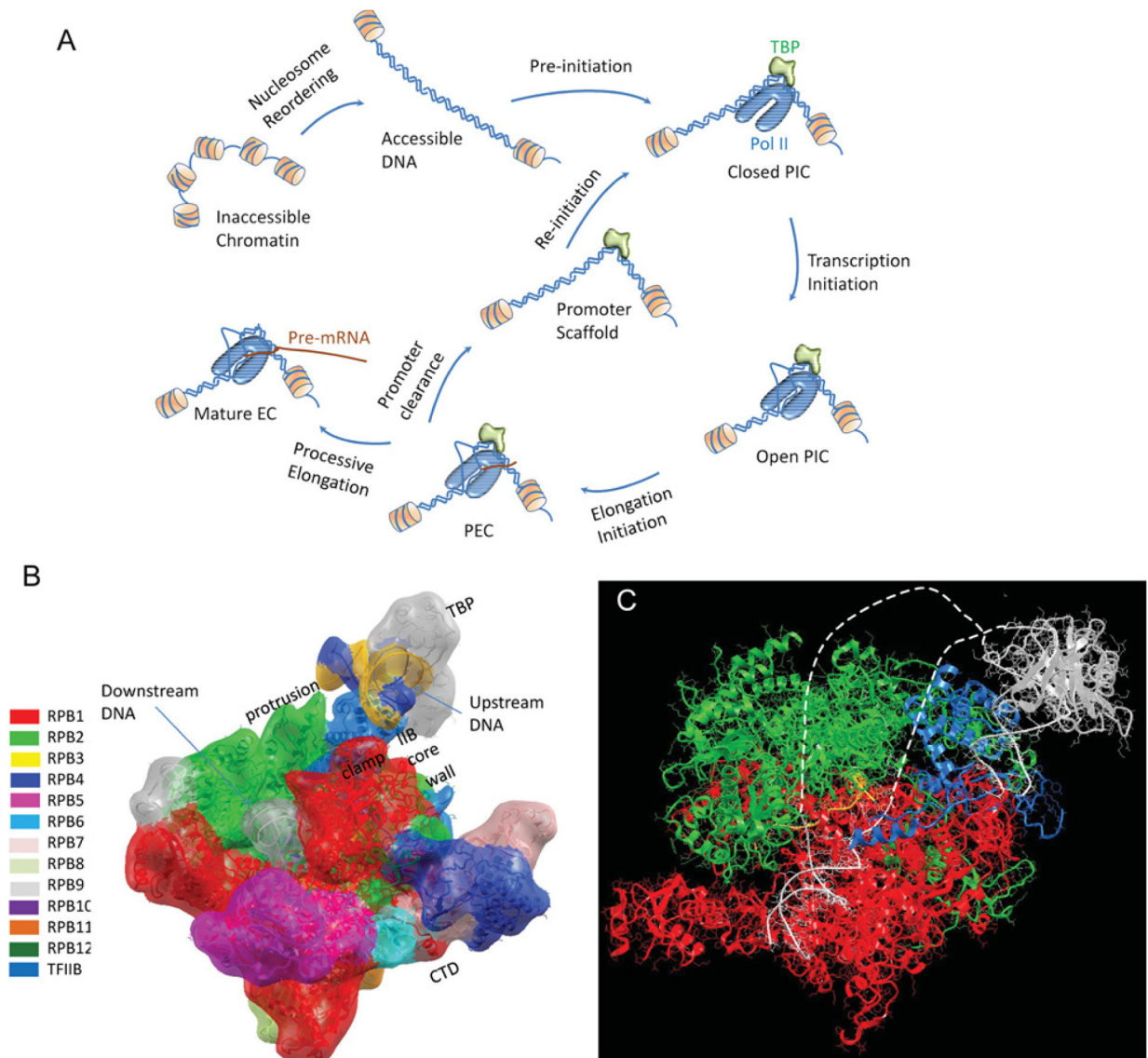
62. Guenther MG, Levine SS, Boyer LA, Jaenisch R and Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130, 77–88 [PubMed: 17632057]
63. Kim TH, Barrera LO, Zheng M, Qu CX, Singer MA, Richmond TA, Wu YN, Green RD and Ren B (2005) A high-resolution map of active promoters in the human genome. *Nature* 436, 876–880 [PubMed: 15988478]
64. Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J and Adelman K (2007) RNA polymerase is poised for activation across the genome. *Nat. Genet* 39, 1507–1511 [PubMed: 17994021]
65. Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M and Young RA (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet* 39, 1512–1516 [PubMed: 17994019]
66. Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA and Young RA (2010) c-Myc regulates transcriptional pause release. *Cell* 141, 432–445 [PubMed: 20434984]
67. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao YA, Li LP and Adelman K (2010) Pausing of RNA Polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143, 540–551 [PubMed: 21074046]
68. Tsai CJ, Ma BY, Sham YY, Kumar S and Nussinov R (2001) Structured disorder and conformational selection. *Protein Struct. Funct. Genet. Bioinfo* 44, 418–427
69. Dyson HJ and Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol* 6, 197–208 [PubMed: 15738986]
70. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D and Nussinov R (2003) Extended disordered proteins: targeting function with less scaffold. *Trends Biochem. Sci* 28, 81–85 [PubMed: 12575995]
71. Gunasekaran K, Tsai CJ and Nussinov R (2004) Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. *J. Mol. Biol* 341, 1327–1341 [PubMed: 15321724]
72. Tompa P and Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem. Sci* 33, 2–8 [PubMed: 18054235]
73. Hegyi H, Schad E and Tompa P (2007) Structural disorder promotes assembly of protein complexes. *BMC Struct. Biol* 7, 65 [PubMed: 17922903]
74. Belakavadi M, Pandey PK, Vijayvargia R and Fondell JD (2008) MED1 phosphorylation promotes its association with Mediator: implications for nuclear receptor signaling. *Mol. Cell. Biol* 28, 3932–3942 [PubMed: 18391015]
75. Asturias FJ, Jiang YW, Myers LC, Gustafsson CM and Kornberg RD (1999) Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* 283, 985–987 [PubMed: 9974391]
76. Dosztanyi Z, Chen J, Dunker AK, Simon I and Tompa P (2006) Disorder and sequence repeats in hub proteins and their implications for network evolution. *J. Proteome Res* 5, 2985–2995 [PubMed: 17081050]
78. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z and Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol* 323, 573–584 [PubMed: 12381310]
79. Taatjes DJ (2010) The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends Biochem. Sci* 35, 315–322 [PubMed: 20299225]
80. Muller F, Demeny MA and Tora L (2007) New problems in RNA Polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. *J. Biol. Chem* 282, 14685–14689 [PubMed: 17395580]
81. Papai G, Tripathi MK, Ruhlmann C, Layer JH, Weil PA and Schultz P (2010) TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* 465, 956–960 [PubMed: 20559389]
82. Papai G, Weil PA and Schultz P (2011) New insights into the function of transcription factor TFIID from recent structural studies. *Curr. Opin. Genet. Dev* 21, 219–224 [PubMed: 21420851]
83. Lyakhov IG, Krishnamachari A and Schneider TD (2008) Discovery of novel tumor suppressor p53 response elements using information theory. *Nucleic Acids Res* 36, 3828–3833 [PubMed: 18495754]

84. Ahn J, Byeon IJL, Byeon CH and Gronenborn AM (2009) Insight into the structural basis of pro- and antiapoptotic p53 Modulation by ASPP proteins. *J. Biol. Chem* 284, 13812–13822 [PubMed: 19246451]
85. Chen WW, Rogatsky I and Garabedian MJ (2006) MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. *Mol. Endocrinol* 20, 560–572 [PubMed: 16239257]
86. Chen W and Roeder RG (2007) The mediator subunit MED1/TRAP220 is required for optimal glucocorticoid receptor-mediated transcription activation. *Nucleic Acids Res* 35, 6161–6169 [PubMed: 17827210]
87. Taubert S, Van Gilst MR, Hansen M and Yamamoto KR (2006) A Mediator subunit, MDT-15, integrates regulation of fatty acid metabolism by NHR-49-dependent and -independent pathways in *C. elegans*. *Genes Dev.* 20, 1137–1149 [PubMed: 16651656]
88. Ito M, Yuan CX, Malik S, Gu W, Fondell JD, Yamamura S, Fu ZY, Zhang XL, Qin J and Roeder RG (1999) Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol. Cell* 3, 361–370 [PubMed: 10198638]
89. Park JM, Kim HS, Han SJ, Hwang MS, Lee YC and Kim YJ (2000) In vivo requirement of activator-specific binding targets of mediator. *Mol. Cell. Biol* 20, 8709–8719 [PubMed: 11073972]
90. Ma BY, Shatsky M, Wolfson HJ and Nussinov R (2002) Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* 11, 184–197 [PubMed: 11790828]
91. Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M and Cramer P (2009) RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* 462, 323–330 [PubMed: 19820686]
92. Liu X, Bushnell DA, Wang D, Calero G and Kornberg RD (2010) Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism. *Science* 327, 206–209 [PubMed: 19965383]
93. Cowling VH (2010) Regulation of mRNA cap methylation. *Biochem. J* 425, 295–302
94. Core LJ, Waterfall JJ and Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848 [PubMed: 19056941]
95. Fuda NJ, Ardehali MB and Lis JT (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186–192 [PubMed: 19741698]
96. Sim RJ, Belotserkovskaya R and Reinberg D (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* 18, 2437–2468 [PubMed: 15489290]
97. Martinez-Rucobo FW, Sainsbury S, Cheung ACM and Cramer P (2011) Architecture of the RNA polymerase–Spt4/5 complex and basis of universal transcription processivity. *EMBO J.* 30, 1302–1310 [PubMed: 21386817]
98. Peterlin BM and Price DH (2006) Controlling the elongation phase of transcription with p-TEFb. *Mol. Cell* 23, 297–305 [PubMed: 16885020]
99. Guiguen A, Soutourina J, Dewez M, Tafforeau L, Dieu M, Raes M, Vandehaute J, Werner M and Hermand D (2007) Recruitment of p-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast. *EMBO J.* 26, 1552–1559 [PubMed: 17332744]
100. Koschubs T, Lorenzen K, Baumli S, Sandstrom S, Heck AJR and Cramer P (2010) Preparation and topology of the Mediator middle module. *Nucleic Acids Res.* 38, 3186–3195 [PubMed: 20123732]
101. D’Alessio JA, Wright KJ and Tjian R (2009) Shifting players and paradigms in cell-specific transcription. *Mol. Cell* 36, 924–931 [PubMed: 20064459]
102. Kenakin T and Miller LJ (2010) Seven transmembrane receptors as shapeshifting proteins: the impact of allosteric modulation and functional selectivity on new drug discovery. *Pharmacol. Rev* 62, 265–304 [PubMed: 20392808]
103. Li H, Gade P, Nallar SC, Raha A, Roy SK, Karra S, Reddy JK, Reddy SP and Kalvakolanu DV (2008) The Med1 subunit of transcriptional mediator plays a central role in regulating CCAAT/enhancer-binding protein- $\beta$ -driven transcription in response to interferon- $\gamma$ . *J. Biol. Chem* 283, 13077–13086 [PubMed: 18339625]

104. Wallberg AE, Yamamura S, Malik S, Spiegelman BM and Roeder RG (2003) Coordination of p300-mediated chromatin remodeling and TRAP/mediator function through coactivator PGC-1  $\alpha$ . *Mol. Cell* 12, 1137–1149 [PubMed: 14636573]
105. Kang YK, Guermah M, Yuan CX and Roeder RG (2002) The TRAP/mediator coactivator complex interacts directly with estrogen receptors  $\alpha$  and  $\beta$  through the TRAP220 subunit and directly enhances estrogen receptor function in vitro. *Proc. Natl. Acad. Sci. U.S.A* 99, 2642–2647 [PubMed: 11867769]
106. Egloff S and Murphy S (2008) Cracking the RNA polymerase IICTD code. *Trends Genet.* 24, 280–288 [PubMed: 18457900]
107. Boeing S, Rigault C, Heidemann M, Eick D and Meisterernst M (2010) RNA polymerase II C-terminal heptarepeat domain ser-7 phosphorylation is established in a Mediator-dependent fashion. *J. Biol. Chem* 285, 188–196 [PubMed: 19901026]
108. Bernecky C, Grob P, Ebmeier CC, Nogales E and Taatjes DJ (2011) Molecular architecture of the human Mediator–RNA polymerase II–TFIIF assembly. *PLoS Biol.* 9, e1000603 [PubMed: 21468301]
109. Swanson MJ, Qiu HF, Sumibcay L, Krueger A, Kim SJ, Natarajan K, Yoon S and Hinnebusch AG (2003) A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol. Cell. Biol* 23, 2800–2820 [PubMed: 12665580]
110. Lee HK, Park UH, Kim EJ and Um SJ (2007) MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* 26, 3545–3557 [PubMed: 17641689]
111. Malik S, Wallberg AE, Kang YK and Roeder RG (2002) TRAP/SMCC/mediator-dependent transcriptional activation from DNA and chromatin templates by orphan nuclear receptor hepatocyte nuclear factor 4. *Mol. Cell. Biol* 22, 5626–5637 [PubMed: 12101254]
112. Malik S, Guermah M, Yuan CX, Wu WZ, Yamamura S and Roeder RG (2004) Structural and functional organization of TRAP220, the TRAP/Mediator subunit that is targeted by nuclear receptors. *Mol. Cell. Biol* 24, 8244–8254 [PubMed: 15340084]
113. Stumpf M, Waskow C, Kroetschel M, van Essen D, Rodriguez P, Zhang X, Guyot B, Roeder RG and Borggrefe T (2006) The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proc. Natl. Acad. Sci. U.S.A* 103, 18504–18509 [PubMed: 17132730]
114. Ge K, Cho YW, Guo H, Hong TB, Guermah M, Ito M, Yu H, Kalkum M and Roeder RG (2008) Alternative mechanisms by which mediator subunit MED1/TRAP220 regulates peroxisome proliferator-activated receptor  $\gamma$ -stimulated adipogenesis and target gene expression. *Mol. Cell. Biol* 28, 1081–1091 [PubMed: 18039840]
115. Udayakumar TS, Belakavadi M, Choi KH, Pandey PK and Fondell JD (2006) Regulation of Aurora-A kinase gene expression via GABP recruitment of TRAP220/MED1. *J. Biol. Chem* 281, 14691–14699 [PubMed: 16574658]
116. Gwack Y, Baek HJ, Nakamura H, Lee SH, Meisterernst M, Roeder RG and Jung JU (2003) Principal role of TRAP/mediator and SWI/SNF complexes in Kaposi's sarcoma-associated herpesvirus RTA-mediated lytic reactivation. *Mol. Cell. Biol* 23, 2055–2067 [PubMed: 12612078]
117. Kim S, Xu X, Hecht A and Boyer TG (2006) Mediator is a transducer of Wnt/ $\beta$ -catenin signaling. *J. Biol. Chem* 281, 14066–14075 [PubMed: 16565090]
118. Zhou HY, Kim S, Ishii S and Boyer TG (2006) Mediator modulates Gli3-dependent sonic hedgehog signaling. *Mol. Cell. Biol* 26, 8667–8682 [PubMed: 17000779]
119. Rau MJ, Fischer S and Neumann CJ (2006) Zebrafish Trap230/Med12 is required as a coactivator for Sox.9-dependent neural crest, cartilage and ear development. *Dev. Biol* 296, 83–93 [PubMed: 16712834]
120. Hallberg M, Hu GZ, Tronnorsjo S, Shaikhibrahim Z, Balciunas D, Bjorklund S and Ronne H (2006) Functional and physical interactions within the middle domain of the yeast mediator. *Mol. Genet. Genom* 276, 197–210

121. Zhang F, Sumibcay L, Hinnebusch AG and Swanson MJ (2004) A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol. Cell. Biol* 24, 6871–6886 [PubMed: 15254252]
122. Natarajan K, Jackson BM, Zhou H, Winston F and Hinnebusch AG (1999) Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator. *Mol. Cell* 4, 657–664 [PubMed: 10549298]
123. Toth JI, Datta S, Athanikar JN, Freedman LP and Osborne TF (2004) Selective coactivator interactions in gene activation by SREBP-1a and -1c. *Mol. Cell. Biol* 24, 8288–8300 [PubMed: 15340088]
124. Kato Y, Habas R, Katsuyama Y, Naar AM and He X (2002) A component of the ARC/mediator complex required for TGF  $\beta$ /Nodal signalling. *Nature* 418, 641–646 [PubMed: 12167862]
125. Yang FJ, Vought BW, Satterlee JS, Walker AK, Sun ZYJ, Watts JL, DeBeaumont R, Saito RM, Hyberts SG, Yang S et al. (2006) An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* 442, 700–704 [PubMed: 16799563]
126. Stevens JL, Cantin GT, Wang G, Shevchenko A and Berk AJ (2002) Transcription control by E1A and MAP kinase pathway via Sur2 mediator subunit. *Science* 296, 755–758 [PubMed: 11934987]
127. Mo XM, Kowenz-Leutz E, Xu H and Leutz A (2004) Ras induces mediator complex exchange on C/EBP  $\beta$ . *Mol. Cell* 13, 241–250 [PubMed: 14759369]

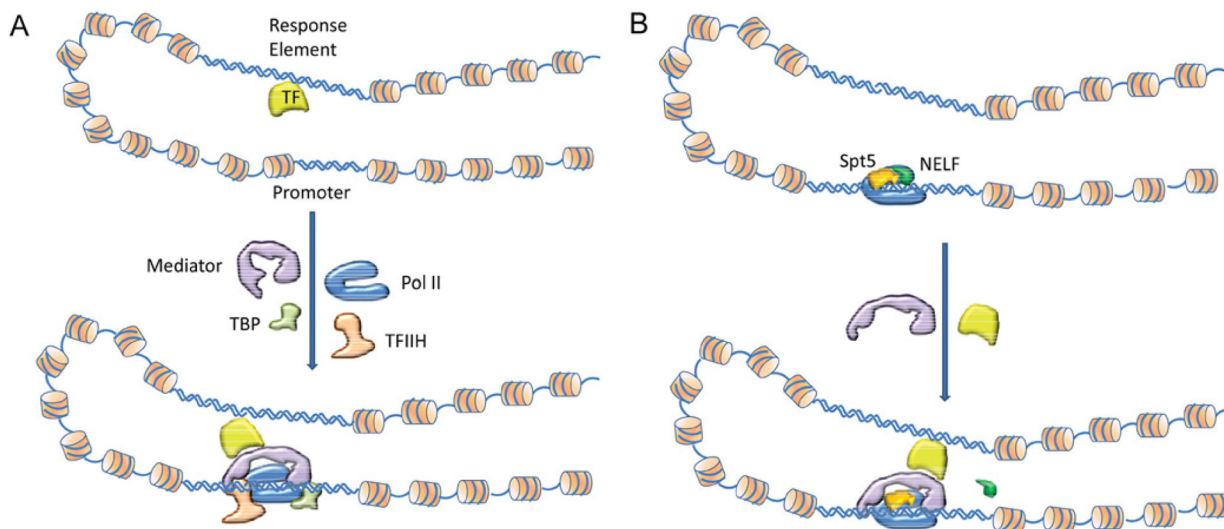




**Figure 1. Simplified illustrations depict distinct stages of RNA Pol II transcription**

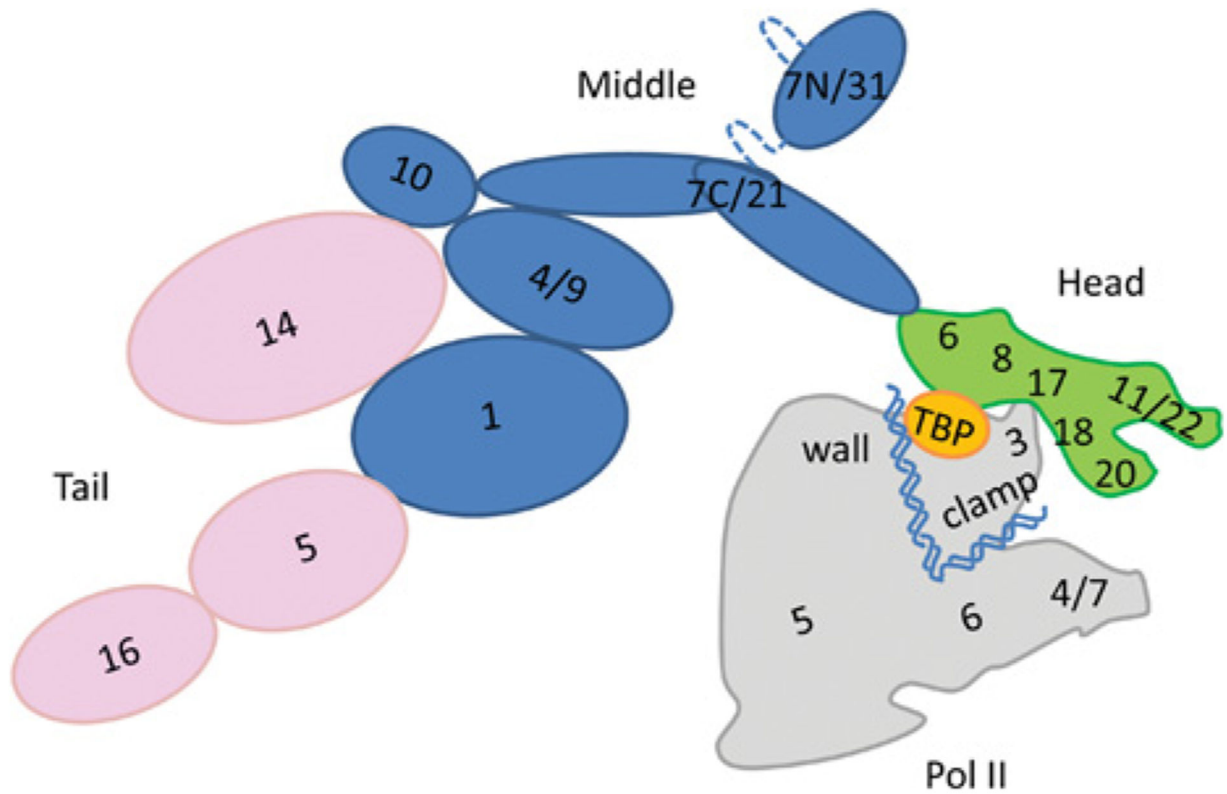
(A) A possible nucleosome reordering process to make the unavailable promoter accessible. This is followed by PIC formation at the promoter region either via an activated recruitment or basal transcription accumulation. EC, elongation complex. (B) Atomic model of the PIC. The relative orientation between the TBP-bound DNA and Pol II is based on their interactions with TFIIB. The signature of transcription initiation is the formation of a transcription bubble with the DNA duplex melting in the Pol II active-site cleft. (C) Atomic model of the open PIC with the melted DNA drawn as broken lines. The open PIC complex then enters the elongation initiation stage to synthesize the first RNA nucleotide. The PEC continues to synthesize the nascent RNA and at a certain RNA length it dissociates from the promoter and leaves the scaffold for another quick re-initiation. At this stage, a mature EC completes the transcription, or a paused EC (not shown in the Figure) stalls at a proximal promoter region, waiting to be activated to resume transcription elongation. Pol II nomenclature is provided in [91,92].



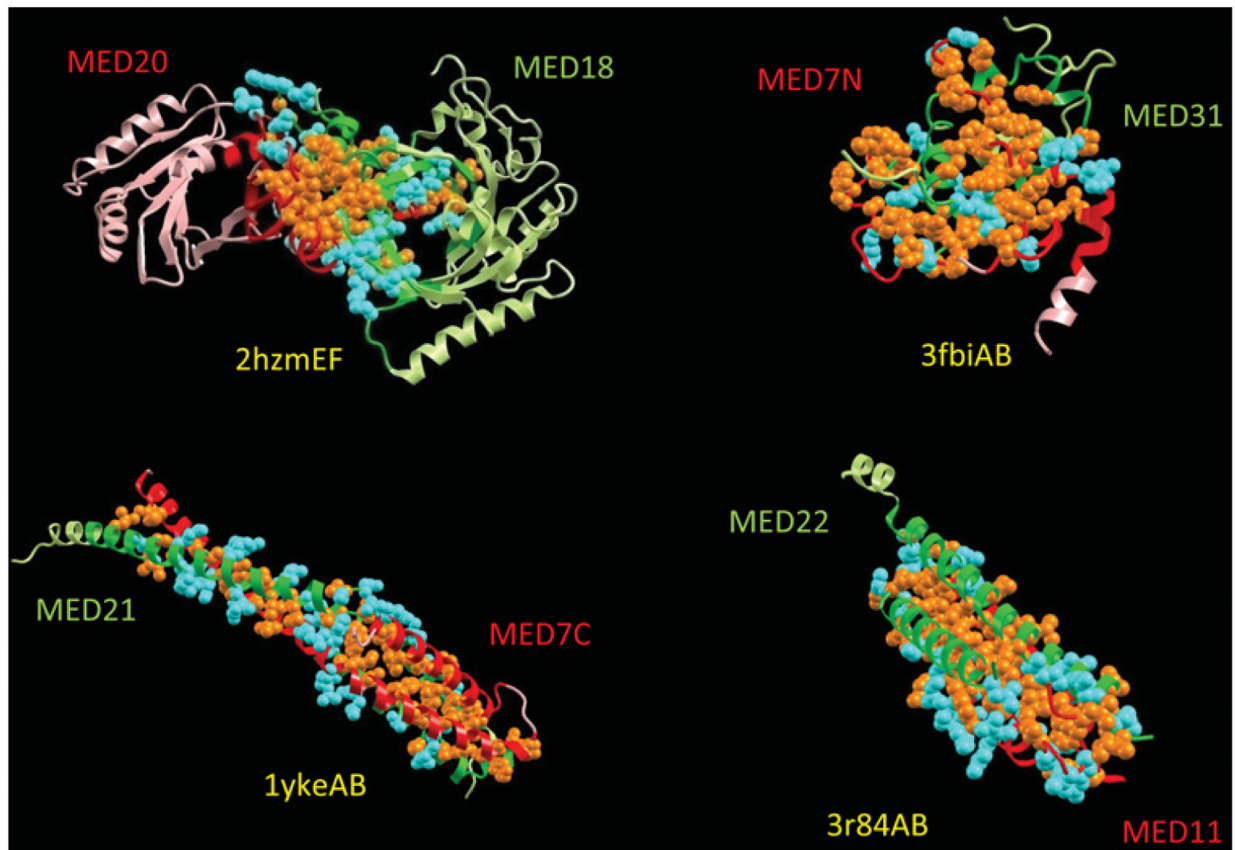


**Figure 2. Schematic diagram illustrating two distinct mechanisms of transcription signal transduction from TF via Mediator to the transcription machinery**

In mechanism I (A), the DNA-bound TF binds Mediator which in turn recruits GTFs/Pol II to complete transcription without interruption. In mechanism II (B), a stalled PEC resumes its processive elongation from a paused stage after the DNA-bound TF binds with Mediator. We assume that the RE is exposed and available for TF binding in both mechanisms, and that the core promoter may be available (nucleosome free) or unavailable for access in mechanism I but is always available in mechanism II. For simplicity and clarity only two GTFs (TFIIH and TBP), Pol II, Mediator, Spt5, NELF and TF are included in the drawing instead of all PEC members. Mechanism I denotes the concept that, following a binding event of the TF DNA-binding domain to the RE, the TF AD recruits Mediator through binding at the Tail/Middle modules. In contrast, in mechanism II PEC already occupies the promoter. After TF binds the RE, it binds to Mediator at the Head/Middle modules and the transcription signal which originates from the TF–RE interaction propagates through Mediator’s Head to Pol II to disassociate NELF and enter processive elongation.



**Figure 3. Schematic two-dimensional organization map of the yeast Mediator and Pol II**  
 The numbers relate to Mediator and Pol II subunits. The broken lines are disordered linkers. This organization is based on experimental data from crystal structures (subunits MED8–MED18–MED20 [36,37], MED7–MED31 and MED7–MED21 [38]); EM (Head module, [25,27]), the Middle and Tail modules have been described in previous studies [41,100]. The relative orientation between Mediator and Pol II has been described previously [108]. Pol II Rpb4/Rpb7 and Clamp bind the Mediator Head module and the Pol II CTD binds the Middle module. In mechanism I, PEC formation is triggered by a TF AD-bound Mediator which interacts with Pol II (red) mostly via its Tail (pink) or Middle (blue) modules. In mechanism II, PEC is pre-assembled on the promoter and consequently there is a low level of Ser<sup>5</sup> CTD phosphorylation. In both mechanisms, the recruitment of p-TEFb is needed to promote processive elongation after the RNA capping checkpoint. We speculate that the perturbation by AD binding to the Mediator Head module (green) causes a significant conformational change which propagates via MED17, MED18 and MED20 to Rpb4/Rpb7 and finally to the Clamp of Pol II. This propagation displaces NELF and facilitates the transition of a stalled PEC into an elongation phase. A large conformational change displayed by a sufficiently high population of Mediator will present ‘shape-shifting’. This map is for yeast; however, the human map is similar. The cartoon does not reflect the actual size of each of the Mediator subunits.



**Figure 4. Protein–protein interfaces of Mediator complexes whose X-ray structures are available** Mediator is a very large multi-protein complex composed of 10–25 subunits. High-resolution structural information on Mediator is extremely limited to small domain–domain interactions or small subunits, all of which constitutes only tiny fractions of a large complex. Protein complexes are represented as ribbons in two light colours with the interface highlighted in dark colours. Only interfacial side chains are shown in space-fill with hydrophobic residues coloured orange and hydrophilic residues in cyan. All three interfaces, MED18/20 (PDB code 2HZM), MED7C/21 (PDB code 1YKE), MED7N/31 (PDB code 3FBI), and MED11/22 (PDB code 3R84), clearly show the domination of hydrophobic interactions which reflect the characteristics of a two-state protein folding. In turn, this explains the specificity and the disorder nature of individual Mediator subunits. Folding-upon-binding transitions of disordered states lead to specific interactions and efficient signal propagation across interfaces, which is particularly important for large multi-chain complexes. Efficient signalling may explain why Mediator subunits are highly disordered. This Figure is available as an interactive three-dimensional structure at <http://www.BiochemJ.org/bj/439/0015/bj4390015add.htm>.

**Table 1**

Mediator modules, subunits and interacting TFs and Pol II CTD

Mediator Module	Subunit	Interacting TF	Reference(s)
Head	MED22	Gcn4	[109]
Head	MED17	VP16	[88]
Head	MED17	PGC-1 $\alpha$	[104]
Head	MED17	p53	[27]
Middle	MED1	p53	[27]
Middle	MED1	C/EBP $\beta$	[103]
Middle	MED1	PGC-1 $\alpha$	[104]
Middle	MED1	ER $\alpha$ ; ER $\beta$	[105]
Middle	MED1	RAR	[110]
Middle	MED1	HNF-4	[111]
Middle	MED1	TR	[112]
Middle	MED1	Pit-1, GATA-1, GATA-2	[113]
Middle	MED1	PPAR $\gamma$	[114]
Middle	MED1	GR	[85,86]
Middle	MED1	GABP	[115]
Middle	MED12	RTA	[116]
Middle	MED12	$\beta$ -catenin	[117]
Middle	MED12	Gli3	[118]
Middle	MED12	Sox9	[119]
Middle	MED21	Tup1	[120]
Tail	MED2	Gcn4	[121,122]
Tail	MED3	Gcn4	[121]
Tail	MED14	HNF-4	[111]
Tail	MED14	GR	[85]
Tail	MED14	SREBP1a	[123]
Tail	MED15	VP16, Gal4	[89]
Tail	MED15	GR, NHP49	[87]
Tail	MED15	Smad2-Smad4	[124]
Tail	MED15	Gcn4	[89,109,121]
Tail	MED15	SREBP1a	[125]
Tail	MED16	PGC-1 $\alpha$	[104]
Tail	MED16	Gcn4	[109]
Tail	MED23	E1A, E1F3	[126]
Tail	MED23	C/EBP $\beta$	[127]
Tail	MED25	VP16	[89]
Tail	MED25	RAR	[110]

Mediator is a huge multi-protein complex (26 subunits in the human [79]). It is the primary regulator of the PEC, which includes TFIIA, TFIIB, TFIID, TFIIIE, TFIIIF, TFIIF, TFIIF, Mediator and Pol II [27,79]. Mediator is divided into Head, Middle and Tail modules. The interactions and functions of some subunits have been partly elucidated [24–27,79]. The Table lists some of the known TF AD interactions [24]. Because only scant structural data are available, we have not distinguished between yeast and human. The CDK8 submodule triggers Mediator conformational change [79],

MED5 has acetyltransferase activity, and Tail MED14 was shown recently to bind the N-terminal domain of PPAR  $\gamma$  [26]. The partial list in this Table illustrates that a TF can bind multiple Mediator subunits even belonging to different modules (e.g. p53, GCN4); however, a TF can also require different subunits, as in p53 activation of p21 by Nutlin which does not require CDK8, whereas UV light C does [60]. Such cases could relate to communication pathways. Secondly, TFs do not necessarily require the same subunits. For example, while MED1 is not required for PPAR  $\gamma$  [26], it is required by, for example, C/EBP $\beta$  [103], PGC-1 $\alpha$  [104] and ER $\alpha$  [105]. Yeast and human subunit interaction maps have been constructed [35,40,41]. The Pol II CTD has up to 52 heptapeptide repeats (YSPTSPS) [106]. Ser<sup>5</sup> and Ser<sup>7</sup> are phosphorylated by CDK7 (TFIIH) which is recruited (with TFIIIB) by Mediator–Pol II complex before transcription initiation [107]. Ser<sup>2</sup> phosphorylation by CDK9 is required for elongation. Ser<sup>5</sup> phosphorylation is required for initiation of transcription elongation. Submodule CDK8 phosphorylates CDK7. Key mechanistic questions relate to how TF–Mediator binding activates Pol II and how this relates to the CTD phosphorylation requirement. C/EBP $\beta$ , CCAAT/enhancer-binding protein  $\beta$ ; E1A, early region 1A; ER, oestrogen receptor; GABP, GA-binding protein; Gli3, GLI family zinc finger 3; GR, glucocorticoid receptor; HNF-4, hepatocyte nuclear factor 4; PGC-1 $\alpha$ , PPAR  $\gamma$  co-activator-1 $\alpha$ ; RTA, replication and transcription activator; Sox9, SRY (sex-determining region Y)-box 9; RAR, retinoic acid receptor; TR, thyroid hormone receptor.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript