Models behind the mystery of establishing enhancer-promoter interactions

Kathryn Monfils, Tahsin Stefan Barakat *

Department of Clinical Genetics, Erasmus MC University Medical Center, PO Box 2040, 3000 CA, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
- Enhancer promoter interaction
- DNA loop
- Transcription initiation
- Condensate

ABSTRACT

Enhancers and promoters are transcriptional regulatory elements whose facilitated interactions increase gene expression. Enhancer DNA sequences can be located far away from the promoter sequences that they regulate. Currently, the mechanism facilitating the establishment of enhancer-promoter interactions remains unclear. However, mutations causing errors in these interactions have been linked to cancer and disease, further conveying the need to understand the full mechanism. This review discusses multiple models that have been proposed to describe how enhancers go the distance to interact with promoters. Evidence supporting loop formation models is reviewed in addition to more complex hypotheses involving aspects of 3D chromatin organization and phase separation.

1. Introduction

Transcriptional regulatory elements are part of the network that controls transcription initiation, and ultimately gene expression. Part of this network is made up of enhancers, or DNA sequences that can control gene expression in a spatiotemporal manner and which can recruit transcription factors (TFs). These factors interact with other components of the transcription machinery bound to gene promoters. Promoters, another class of regulatory elements, are DNA sequences that contain the gene transcription start site. When the interaction between enhancers and promoters takes place in eukaryotes, transcription will be initiated by RNA Polymerase II (Pol II). Based on large-scale epigenome characterization, the ENCODE Project Consortium has predicted that there are millions of candidate enhancers in humans, vastly outweighing the number of protein coding genes (Karnuta and Scacheri, 2018). An increasing number of studies have linked genetic mutations in enhancer elements to disease (Carullo and Day, 2019; Herz, 2016; Perenthaler et al., 2019; Sakabe et al., 2012; Yousefi et al., 2021), and in addition, disturbance of the 3D genome organization which can disrupt interactions between enhancers and promoter has been identified in a number of genetic disorders (Spielmann et al., 2018). Taken together, this explains why fundamental research remains crucial to understand the molecular mechanisms underlying enhancer-promoter (E–P) interactions.

Enhancers and promoters are active within topologically associated domains (TADs) of chromosomes, which are sub-megabase, isolated regions in which chromatin interactions preferably take place; this provides an additional layer of regulation. In order to increase transcription rates, enhancers and promoters need to be in physical proximity to each other, and possibly even make direct contact. Even though TADs create a smaller playing field for E–P interactions to occur, a challenging mechanistic feat remains, given that enhancers can be located hundreds of kilobases away from the promoters that they regulate. The fact that E–P interactions occur has been long accepted, but currently, there is no consensus regarding precisely how these interactions are established.

In this mini-review, we discuss various models that have been suggested to explain the establishment of E–P interactions, as well as the science that supports them. We begin by analyzing proposed models that involve chromatin looping: tracking, DNA looping, and loop extrusion. Then, we continue with more all-encompassing models such as the ‘selecting-facilitating-specifying’ model and the transcription condensates theory. Lastly, we briefly deliberate on whether transcription condensates are causes or consequences of E–P interactions.

2. Loop formation models

Around the turn of the century, researchers had discovered that enhancers control gene expression through long-range interactions, and that the chromatin between enhancers and promoters was physically looped out of the way (reviewed in (Bulger and Groudine, 1999; Carter et al., 2002; Tolhuis et al., 2002). The mechanism resulting in loop formation and specific binding was still unclear, and research began uncovering plausible models to explain this phenomenon.
One of these models is the tracking model, which refers to the mechanism that Pol II uses after binding the enhancer, to move along the DNA in order to bind the promoter (Fig. 1A) (Blackwood and Kadonaga, 1998; Hatzis and Talianidis, 2002; Wang et al., 2005). Assuming that Pol II and other enhancer-bound activators travel along the DNA strand in a unidirectional manner, once contact has been achieved with the promoter, the DNA strand has successfully formed a loop structure. Supporting this model are chromatin immunoprecipitation results that have detected unidirectional movements of the DNA-TF-enhancer complex on DNA moving towards the promoter (Hatzis and Talianidis, 2002). Other evidence is the existence of intergenic RNAs, including short, polyadenylated RNAs transcribed from the enhancer DNA, which sometimes span the region between the enhancer and the promoter (Kong et al., 1997; Tchurikov et al., 2009; Zhu et al., 2007). This tracking mechanism was tested by the insertion of DNA insulator elements between enhancers and promoters, which appear to halt the movement of the tracking proteins at the inserted insulator elements, subsequently preventing E–P interactions (Tchurikov et al., 2009; Zhu et al., 2007). However, other studies have concluded that insulators block E–P communication without blocking transcription through the insulator (Bender and Fitzgerald, 2002; Fujioka et al., 2021). Furthermore, the tracking model cannot explain how enhancers and promoters interact inter-chromosomally (Lomvardas et al., 2006; Zhao et al., 2006), which suggests that E–P interactions occur either by different or by multiple mechanisms.

Partially due to the long distances involved, another accepted model behind E–P interactions is the DNA looping model, wherein the dynamic binding of enhancer-bound TFs with the TFs bound to the promoter connect the two directly (Fig. 1B) (Carter et al., 2002; Cook and Marenduzzo, 2018; de Laat and Grosveld, 2003; Jhunjhunwala et al., 2008; Ptashne and Gann, 1997). If the two elements are in cis, as this contact is made, the DNA in between is looped out, (Tolhuis et al., 2002). If located on different chromosomes, such contacts could still occur, but no looping would be involved. Supporting the TF looping mechanism, research shows chromatin looping can be forced by using zinc finger DNA binding proteins to establish distant E–P interactions (Bartman et al., 2016; Deng et al., 2012, 2014). This forced chromatin looping technique was able to initiate β-globin transcription even in the absence of GATA1, a specific TF required for chromatin looping at the β-globin locus (Deng et al., 2012). Additionally, forced looping was able to trigger transcriptional reactivation of developmentally silenced embryonic globin genes in adult murine erythroblasts (Deng et al., 2014). These results suggest that the sole role of some key TFs is to establish direct E–P contacts, which consequently results in gene expression.

While dynamic looping of TFs seems plausible, it has also been suggested that a combination of looping and tracking may establish E–P interactions. For instance, in many cases, the dynamic binding of TFs to each other may be inefficient if there was not some kind of facilitation of loop formation. Direct TF contacts may only guide the regulatory element-bound TFs to the same general vicinity, and a scanning or tracking mechanism may finish the job. While we know much about the consequences of these interactions, the driving factors or recruitment signals behind these models are still not completely understood.

The loop extrusion model, which underlies the formation of TADs, may give more insight into how E–P interactions can be established in an efficient manner. This model proposes that loop extruding factors, such as the cohesin protein complex, hold two adjacent regions of DNA while at the same time extruding a loop by translocating along the chromatin.

---

**Fig. 1.** Schematic illustrating DNA tracking, DNA looping, and loop extrusion models. **A.** DNA tracking: TFs bind to the enhancer and track along the DNA in a unidirectional manner until reaching the promoter. **B.** DNA looping: TFs dynamically bind to each other, directly connecting the enhancer and the promoter. Both cis and trans interactions are pictured. **C.** Loop extrusion: the cohesin complex extrudes DNA, creating a loop configuration. Extrusion halts at CTCF binding sites which are located close to transcriptional regulatory elements or insulators. (Figure created using Biorender.com.)
fiber in one or both directions (Davidson et al., 2019; Fudenberg et al., 2016; Kim et al., 2019). It is hypothesized that such a complex may stall at an insulator element, keeping an active regulatory element in place on one side, while continuing to extrude DNA on the other – essentially searching for another insulator element and neighboring regulatory element (Fig. 1C) (Fudenberg et al., 2016). The fact that binding sites for cohesin and CCCTC-binding factor (CTCF) are found near enhancers and promoters, in addition to those found at the boundaries of TADs, support this concept (Dowen et al., 2014; Ing-Simmons et al., 2015; Stadhouders et al., 2019). Moreover, the highly sensitive Micro-Capture-C method links the activation of promoters and enhancers to increased loading of cohesin (Hua et al., 2021). Aside from this evidence, it is easy to draw similarities between the tracking model and the loop extrusion hypothesis. However, what causes the cohesin complex to stall at insulator elements has not yet been identified. Although loop extrusion is now widely accepted, the precise details and order of events remain to be determined. Furthermore, there are some observations that do not fully support this as the only mechanism, examples being that E–P contacts can be cohesin-independent and that cohesin depletion does not substantially decrease gene expression, although significant changes do occur (Andrey et al., 2017; Monahan et al., 2019; Rao et al., 2014; Schwarzer et al., 2017; Thiecke et al., 2020). Additionally, while the Micro-Capture-C method does support some elements of this model, results from the same study also support the maintenance of E–P interactions via DNA-binding proteins (Hua et al., 2021). This suggests that E–P interactions do not all occur via one mechanism, but rather that the interactions are a result of a complex combination of mechanisms.

3. Multi-layered models

In 2019, a ‘selecting-facilitating-specifying’ model was proposed by Schoenfelder and Fraser which we discuss here as a valid contender to explain the multiple facets of E–P interaction establishment (Schoenfelder and Fraser, 2019). Simply put, the model consists of three categories, each of which is essential for productive E–P interactions.

First, genome sequences are ‘selected’ and modified by TFs and chromatin modifying enzymes in order to mark the sequences as regulatory elements. For example, pioneer TFs, which are the first TFs to engage with compact chromatin, pave the way for other TFs by opening up chromatin, allowing access (reviewed in (Zaret and Carroll, 2011)). Additionally, active promoters and enhancers are marked with the histone modifications H3K4me3 and H3K27ac, respectively (reviewed in (Kang et al., 2020)). These changes are the first steps for creating space and for marking the regulatory elements for future contacts.

Secondly, the regulatory elements are ‘facilitated’ into environments that will reduce the search space (e.g., active and inactive chromatin compartments, TADs, extrusion loops), thereby increasing the encounter chance between enhancers and promoters. A study has shown that changing distances between regulatory elements located within a TAD has relatively little effect on gene expression (Symmons et al., 2016). Meanwhile, disruption of TAD formation can cause loss of gene expression, accompanied by reduction of E–P contacts or loss of contacts all together (Symmons et al., 2016). Confoundingly, other work suggests that boundary elements of TADs are not always required for high-level gene expression. Using transgene constructs flanked by TAD boundary elements, TAD formation was mimicked, and transcription levels were high (Yokoshi et al., 2020). However, when one boundary element was deleted or inverted, TAD formation was no longer detected, yet E–P interactions still occurred. In addition to these results, it is known that enhancers occasionally interact outside of their designated TADs (Franke et al., 2016). This would suggest that 3D chromatin organization plays a role in E–P interactions, but that it is not the only defining component in their establishment.

The third category is ‘specifying’: TFs and bridging proteins that are bound to regulatory elements interact with each other, bringing the promoter and enhancer into close enough proximity to initiate transcription. This specificity has been previously tested by inserting TF binding sites as decays between enhancers and promoters, which can result in nonproductive, yet specific, looping interactions (Nolls et al., 2009). In line with this model, results show that artificial looping bypasses the need for some TFs, as previously mentioned (Deng et al., 2012, 2014), and that specific TF knockout mutants fail to form proper E–P interactions (Drissen et al., 2004; Vakoc et al., 2005).

The authors of the ‘selecting, facilitating, specifying’ model also include the concept of liquid-liquid phase separation as a means for facilitating the establishment of E–P interactions, as also suggested by others (Cramer, 2019; Di Giammartino et al., 2020; Robson et al., 2019; Schoenfelder and Fraser, 2019). Phase separation occurs when intrinsically disordered regions (IDRs) of proteins, or in this case TFs, engage in weakionic interactions with other IDRs, resulting in condensates, or membrane-less organelles (Banani et al., 2017). In other words, high concentrations of IDRs in a small space allow the proteins to de-emulsify, just as water and oil can separate under the right conditions. Convincingly, transcriptional components such as Mediator, BRD4, and Pol II contain IDRs and have been shown to participate in phase separation in vitro (Boija et al., 2018; Cho et al., 2018; Chong et al., 2018). The proposed model for E–P interactions is therefore that TFs or activator proteins bind to enhancers, and their IDRs ‘recruit’ other IDRs to the same region, establishing a phase condensate (Fig. 2A). This same recruitment would cause promoters to enter the condensate, facilitating E–P interactions that ultimately initiate transcription. Thus, the condensate provides the means for increased E–P interactions without directing specific contacts. While the scientific community is still divided, and the final verdict is not in; condensates could explain previously confusing findings such as enhancers activating multiple promoters simultaneously (Fukaya et al., 2016; Lim et al., 2018) and transcriptional bursting occurring without direct E–P contact (Alexander et al., 2019).

4. Condensates and E–P interactions: cause or coincidence

With all the advances recently made regarding the link between phase separation and transcription, the establishment of E–P interactions is often considered to be a product of the formed condensates. However, to date, there is no concrete evidence directly connecting phase separation with the establishment of E–P interactions. In fact, most of the evidence linking condensates to E–P interactions revolves around super-enhancers, regions that are associated with high concentrations of transcriptional elements and their regulators (reviewed in (Wang et al., 2019)). Therefore, we would like to raise the question whether phase separation is really the cause of E–P interactions or if condensates are formed in parallel to the still not fully defined forces driving E–P interactions.

Building upon the knowledge that condensates form around IDRs in TFs (including the Pol II holoenzyme) that bind to regulatory elements, we can conclude that both enhancers and promoters are surrounded by some pre-condensate form when active. It has been suggested that the coalescence of condensates could drive E–P interactions (Cho et al., 2018; Hu and Tee, 2017). This coalescence was visualized through a novel, targeted, condensation technique called CasDrop. This technique uses CRISPR-Cas9 technology to insert an ‘optogenetic assembly’ that contains a fluorescent marker and binding scaffolds for recruiting proteins with IDRs, among other components (Shin et al., 2018). Results show that genomic loci containing CasDrop condensates can be pulled together to form larger condensates, while non-targeted chromatin is pushed out of the way (Shin et al., 2018). Another study shows that condensates help to shorten the target-search process of transcriptional elements and reduces the number of non-specific trials made (Kent et al., 2020).

Computational results also support a link between long-range interactions and condensate formation by showing that distant TF binding sites are able to join together to form a single condensate (Shrinivas et al., 2019). In this same study, specific DNA-TF interactions
were shown to be required for both the formation and stability of condensates, where they compensate for the loss of free energy in the nucleus that is associated with phase separation. Based on these studies, it seems that condensates associated with TFs provide a setting favorable for long-range, specific E–P interactions.

However, it can also be argued that condensates are merely a product of the situation that generates such E–P interactions. For example, untargeted droplet formation shows that droplets, or condensates, are more frequently found in regions of euchromatin, or so it appears (Shin et al., 2018). A mathematical model predicts that droplets can be formed in both euchromatin and heterochromatin. Although low-density euchromatic regions are more favorable to droplet growth because of the extra space that they encompass, this likely also allows for easy optical detection (Shin et al., 2018). Additionally, phase-separated condensates are only observed when TFs are over-abundant, while lower, endogenous TF expression levels only promote pre-condensate, ‘hub’ formation, and not actual phase separation (Chong et al., 2018). The most convincing evidence to date comes from recently published results showing that by dissolving condensates and releasing the TFs BRD4 and Mediator, through means of the compound 1,6-hexanediol and a small molecule inhibitor, respectively, there is a negative impact on gene transcription while E–P looping structures remain intact (Fig. 2B) (Crump et al., 2021). These results suggest that condensates are not required for the establishment or maintenance of E–P interactions, but can play a parallel, independent role in transcription regulation.

5. Conclusion

The various models discussed here have all been proposed to explain how E–P interactions are established. From simple looping models to complex, multi-level theories, it seems clear that the mechanism behind these interactions has not yet been completely uncovered. Given the sometimes contradictory evidence, the prospects that multiple mechanisms exist at different loci and that short-range and long-range interactions occur differently are still possible. While advances in technology such as high-resolution imaging and chromatin capture techniques are driving progress in this field, we believe that large-scale collaborative efforts will be needed to solve this mystery. Varying aspects of the models presented here need to be tested in standardized temporal experiments on both short- and long-range E–P interactions. Only then will the scientific community finally be able to draw conclusions on exactly how, when, and why E–P interactions are established. Ultimately, this fundamental understanding will have an impact on multiple research fields such as transcriptional cell biology, molecular genetics, cellular mechanics, and cancer research.

Author contributions

KM: Investigation, Writing – Original draft, Visualization.

TSB: Conceptualization, funding acquisition, supervision, Writing – Review & Editing.

Funding

TSB’s lab is supported by the Netherlands Organisation for Scientific Research (ZonMW Veni, grant91617021), a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation, an Erasmus MC Fellowship 2017, and Erasmus MC Human Disease Model Award.
2018.

Declaration of Competing Interest
The authors report no declarations of interest.

Acknowledgements
We would like to thank all members of the Barakat lab for helpful discussion.

References
Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., Weiner, O.D., 2019. Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer activity. Cell 179, 1865–1875.e1816.
Bartman, C.R., Hsu, S.C., Hsiung, C.C.S., Raj, A., Blobel, G.A., 2016. Enhancer regulation of gene expression and diseases. Biosci. Rep. 37, BSR20160183.
Drissen, R., Palstra, R.-J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., de Carullo, N.V.N., Day, J.J., 2019. Genomic enhancers in brain health and disease. Genes Brain Behav. 18, 393–406.
Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Trifonov, E.N., Mirny, L.A., 2016. Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. Genes Dev. 30, 1013–1036.
Karnuta, J.M., Scacheri, P.C., 2018. Enhancers: bridging the gap between gene control and human disease. Hum. Mol. Genet. 27, R219–R227.
Kraft, K., Kempfer, R., Hansen, L.L.P., Hestand, L.D., Smit, M.L., 2012. Enhancer-promoter interactions regulate olfactory receptor choice. Nature 565, 448–453.
Kraft, K., Kempfer, R., Hansen, L.L.P., Hestand, L.D., Smit, M.L., 2012. Enhancer-promoter interactions regulate olfactory receptor choice. Nature 565, 448–453.
Kraft, K., Kempfer, R., Hansen, L.L.P., Hestand, L.D., Smit, M.L., 2012. Enhancer-promoter interactions regulate olfactory receptor choice. Nature 565, 448–453.
Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., de Laat, W., 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol. Cell 10, 1453-1465.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., Blobel, G.A., 2005. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol. Cell 17, 453-462.

Wang, Q., Carroll, J.S., Brown, M., 2005. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. Mol. Cell 19, 631-642.

Wang, X., Cairns, M.J., Yan, J., 2019. Super-enhancers in transcriptional regulation and genome organization. Nucleic Acids Res. 47, 11481-11496.

Yokoshi, M., Segawa, K., Fukaya, T., 2020. Visualizing the role of boundary elements in enhancer-promoter communication. Mol. Cell 78, 224–235 e225.

Yousefi, S., Deng, R., Lanko, K., Salsench, E.M., Nikoncuk, A., van der Linde, H.C., Perenthaler, E., van Hau, T., Mulugeta, E., Barakat, T.S., 2021. Comprehensive multi-omics integration identifies differentially active enhancers during human brain development with clinical relevance. bioRxiv 2021, 2004, 2005.438382.

Zaret, K.S., Carroll, J.S., 2011. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 25, 2227-2241.

Zhao, Z., Tavoosidana, G., Sjolander, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S., Ohlsson, R., 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat. Genet. 38, 1341-1347.

Zhu, X., Ling, J., Zhang, L., Pi, W., Wu, M., Tuan, D., 2007. A facilitated tracking and transcription mechanism of long-range enhancer function. Nucleic Acids Res. 35, 5532-5544.