Review

CellPress

Special Issue: Chromatin Dynamics

In search of the determinants of enhancer–promoter interaction specificity

Joris van Arensbergen[1], Bas van Steensel[1], and Harmen J. Bussemaker[2,3]

[1] Division of Gene Regulation, Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands
[2] Department of Biological Sciences, Columbia University, New York, NY 10027, USA
[3] Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10032, USA

Although it was originally believed that enhancers activate only the nearest promoter, recent global analyses enabled by high-throughput technology suggest that the network of enhancer–promoter interactions is far more complex. The mechanisms that determine the specificity of enhancer–promoter interactions are still poorly understood, but they are thought to include biochemical compatibility, constraints imposed by the three-dimensional architecture of chromosomes, insulator elements, and possibly the effects of local chromatin composition. In this review, we assess the current insights into these determinants, and highlight the functional genomic approaches that will lead the way towards better mechanistic understanding.

New approaches to a classic problem

Enhancers have long been known to play a crucial role in orchestrating the genome-wide transcriptional landscape across various cell types, and in response to a broad variety of signals. They are functionally defined as sequence elements that, when linked *in cis* to a promoter, can stimulate its activity, irrespective of orientation [1–3]. Enhancers are typically a few hundred bp long and can harbor binding sites for a wide variety of transcription factors (TFs). Promoters are generally defined as the region immediately surrounding the transcription start site (TSS) at which the transcription pre-initiation complex is assembled, plus the surrounding sequence at which the regulatory input of the gene is integrated [4].

How do enhancers affect promoter activity? In this review, we will focus on the currently dominant view, in which enhancers influence promoter activity through encounters in three-dimensional space. We note, however, that alternative mechanisms have been proposed. For example, enhancers can act as nucleation sites for the establishment of large domains of transcriptionally permissive chromatin, which in turn contributes to the activity of promoters ([5]; for review see [1]).

Although it was originally thought that enhancers regulate only a single nearby promoter, many observations over the past 25 years point to a more complex interplay. Enhancers can control multiple neighboring genes [6–10], sometimes over hundreds of kb and often skipping one or more genes [11,12]. Enhancer–promoter interactions have even been reported to occur between chromosomes [13], but these are rare. As we will discuss below, recent systematic surveys based on high-throughput assays also indicate that the global network of enhancer–promoter interactions may be much more complex than previously believed.

Global mapping of encounters between promoters and enhancers

Physical *in vivo* encounters between pairs of genomic loci in general, and between promoters and enhancers in particular, can be mapped using a family of high-throughput assays named 4C, 5C, and Hi-C, which are based on the core technology of chromatin conformation capture (3C) [14]. All employ *in situ* cross-linking followed by proximity ligation, but differ in the scope and density of coverage of the huge space of potential locus–locus combinations [15]. ChIA-PET is a variant technology that includes an additional immunoprecipitation step that enriches for the presence of a TF of choice [16].

Recent studies using 3C-based techniques have provided initial maps of distal enhancer–promoter contacts. Active promoters were found, on average, to contact 4–5 enhancer-like elements. The majority of these elements are located within 500 kb from the interacting promoter, with an estimated median distance of ~125 kb [17,18]. Interestingly, active enhancers were found to contact approximately two promoters on average, suggesting that enhancers might commonly regulate multiple genes. Moreover, only a fraction of looping distal elements contact the nearest promoter (reported as 27% in [18] and 60% in [19]), whereas the others skip one or more genes. These data indicate that it is often incorrect to assume that an enhancer interacts only with its nearest promoter. Adding to the complexity, a ChIA-PET survey of loci bound by RNA polymerase II found that promoters often interact with

Corresponding authors: van Steensel, B. (b.v.steensel@nki.nl); Bussemaker, H.J. (hjb2004@columbia.edu).
Keywords: gene regulation; enhancer–promoter interaction; biochemical compatibility; chromosome architecture.

0962-8924/
 http://dx.doi.org/10.1016/j.tcb.2014.07.004

other promoters; some of these promoters are able to act as enhancers for their partner promoter, suggesting regulatory cross-talk between promoters resembling enhancer–promoter interactions [19].

Several studies have indicated that the genome-wide pattern of enhancer–promoter interactions is substantially different from one cell type to another [17,18,20,21], although a recent study performing 4C for ~100 *Drosophila* enhancers found that very few changes between mesoderm and whole embryos at two stages [22]. Treatment of human fibroblast cells with TNF-α also left the enhancer-promoter contacts largely unaltered even though ~800 genes showed differential expression [17]. Thus, the network of enhancer–promoter interactions may be more or less reorganized, depending on the cellular state transitions that are studied.

A hierarchy of encounters between enhancers and promoters

Although the 3C-based results discussed above are tantalizing, they come with caveats that closely parallel the complexities associated with the functional interpretation of *in vivo* TF occupancy data. With regard to the latter, it has become evident over the past several years that the relationship between *in vivo* ChIP enrichment and DNA binding affinity of TFs is rather complex [23,24]. Moreover, TF binding does not necessarily lead to regulation of nearby genes [25]. Similar distinctions have to be kept in mind when interpreting the results of 3C-type experiments: a first question is to what extent enhanced frequency of physical contacts between a pair of genomic loci implies a *direct* molecular interaction; and a second question is to what extent detected enhancer–promoter interactions are 'functional', here defined as promoting transcriptional activity (Box 1).

With regard to the first caveat, due to the size of the DNA fragments analyzed (~4 kb on average), elements annotated in these studies as 'promoters' are in fact promoters plus several kb of flanking DNA. Hence, some of the interactions assigned to promoters may, in fact, be driven by flanking DNA elements. Moreover, the spatial resolution of 3C methods, that is, how close two DNA elements need to be in order to be captured by the crosslinking, is still matter of debate. It may well be that a sizeable fraction of the encounters reported by the current methods reflect indirect contacts: for example, because the enhancer and promoter are both part of a much larger structure such as the hypothesized 'transcription factory' in which several genes and their regulatory elements may congregate [26].

Box 1. Three kinds of encounters between enhancers and promoters

Taking inspiration from the title of a review by Palstra [95], we propose the following hierarchy of 'close encounters' between enhancers and promoters:

Close encounters of the* first *kind ('contacts') – Enhancer and promoter show evidence of physical proximity according to assays from the 3C family. Therefore, these encompass all 3C-based identified contacts that occur at a frequency above the (distance-corrected) background, but presumably include many indirect contacts.

Close encounters of the* second *kind ('direct contacts') – Enhancer and promoter interact directly and specifically through molecular recognition mechanisms. Therefore, these constitute a subset of the 'contacts', which is not distinguishable by current 3C-based techniques.

***Close encounters of the* third *kind* ('functional contacts') –** The enhancer and promoter contact each other (directly or indirectly) and this contact has a functional effect on the expression of the gene controlled by the promoter.

Correlative strategies: strengths and limitations

One of the strategies that have been used to identify putative functional interactions between distal regulatory elements and TSSs relied on the mapping DNase hypersensitive sites (DHSs) – generally thought to be active regulatory elements – across 79 different cell types [27]. The correlated presence of DHSs at a promoter locus and distal loci up to 500 kb away was taken as evidence for a functional promoter–enhancer interaction. Many enhancer–promoter pairs showed significant correlation. A modest fraction (4%) of these overlapped with physical interactions identified using 5C or ChIA-PET [27,28]. This is perhaps not surprising, as correlations can arise in multiple ways, including through indirect and noncausal associations.

In an alternative, but conceptually similar approach, yielding similar conclusions, the chromatin state of enhancers was correlated with that of promoters in order to identify functional enhancer–promoter pairs [29]. Like 3C-based assays, these approaches linked enhancers to multiple promoters and *vice versa*. For example, approximately half of the promoter-correlated distal DHSs were assigned to more than one promoter, and approximately half of the promoters were associated to >10 distal DHSs [27]. Finally, a recent extensive atlas of enhancer RNAs (eRNAs) in ~800 different samples from human primary cells, tissues, and cell lines was used to associate enhancers with promoters [28]. Here, on average, promoters were linked to approximately five enhancers, and enhancers were associated with approximately two promoters. Again, in this case, the detected correlations are not guaranteed to correspond to direct causal effects. Indeed, of the inferred functional interactions, 21% were supported by physical contacts based on ChIA-PET analysis [19]. These interactions are the most likely candidates to be biologically relevant.

The molecular mechanisms underlying locus–locus interaction specificity

As described above, both classic and genome-wide studies have highlighted the complex nature of the network of physical and functional associations between enhancers and promoters. While many promoters appear to interact with multiple enhancers, and conversely many enhancers interact with multiple promoters, a remarkable degree of specificity is observed. This raises the question how promoters and enhancers 'choose' each other. Below, we discuss several mechanisms that may underlie this mutual selectivity (Figure 1):

(i) biochemical compatibility;
(ii) spatial architecture of chromosomes within the nucleus;
(iii) insulator elements; and
(iv) local chromatin composition.