Published in final edited form as:

Trends Cell Biol. 2014 November ; 24(11): 695–702. doi:10.1016/j.tcb.2014.07.004.

In search of the determinants of enhancer–promoter interaction specificity

Joris van Arensbergen¹, Bas van Steensel¹, 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

Abstract

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.

Keywords

gene regulation; enhancer–promoter interaction; biochemical compatibility; chromosome architecture

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 vivo* 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 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].

As may be expected, 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 in whole embryo and mesoderm tissue at two developmental stages found that very few (~6%) contacts showed significant changes [22]. Providing insights into the effect of signaling events, it has been shown that treatment of fibroblast cells with TNF- α leaves the enhancer–promoter contacts largely unaltered even though ~800 genes showed differential expression [17]. Therefore, as yet, there is no convergent view on the degree of reshaping of the network of promoter–enhancer

interactions during development, which might reflect differences between lineages and developmental stages 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].

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:

- i. biochemical compatibility;
- ii. spatial architecture of chromosomes within the nucleus;
- iii. insulator elements; and
- iv. local chromatin composition (Figure 1).

We note that, although for practical reasons we conceptualize them as separate mechanisms, often no strict separation exists between them.

Biochemical compatibility

One mechanism that can contribute to the specificity of promoter–enhancer interactions is biochemical compatibility, that is, the intrinsic ability of a promoter–enhancer pair to engage in a specific interaction through recruitment of the right combination of proteins, which in turn is encoded in the DNA sequences of the enhancer and promoter. This concept is illustrated by the classic example of the genes *gsb* and *gsbn* in *Drosophila*, which are two juxtaposed, divergently transcribed genes that have different expression patterns. The enhancers (GsbE and GsbnE, respectively) that drive these different expression patterns are both located in the ~10-kb region that separates their TSSs, and thus could potentially cross-activate the other promoter; however, they do not. Swapping experiments demonstrated that GsbE can only activate the *gsb* promoter, while GsbnE can only activate the *gsbn* promoter [30]. Therefore, these enhancers exhibit distinct promoter compatibilities. Another example of biochemical compatibility is provided by the *dpp* enhancer, which selectively activates *dpp* and not other more proximal lying genes [31].

Later studies in *Drosophila* demonstrated that the core-promoter elements, specifically the TATA-box and the downstream promoter element (DPE), can be critical factors in determining compatibility with certain enhancers. Upstream binding motifs for the TF Caudal showed a preference to activate a DPE-containing promoter rather than a TATA-containing promoter, providing a clue to a possible compatibility mechanism [32]. Another study identified three DPE-specific enhancers, one TATA-specific enhancer, and 14 apparently nonspecific enhancers [33]. It should be noted, however, that in the used enhancer-trap strategy, competition occurs with the endogenous target of the enhancer and,

therefore, the outcome might reflect a preference of an enhancer for a certain promoter-type rather than absolute incompatibility, which may be rare. An example is the IAB5 enhancer of the Bithorax complex. IAB5 preferentially activated a TATA-containing promoter but could activate a DPE-containing promoter if no TATA containing promoter was available [34]. Importantly, however, neither of the genes at the endogenous locus that are proximal to the IAB5 enhancer – *AbdA* and *AbdB*, located at 48 kb and 55 kb, respectively – have a TATA box. The selective activation of *AbdB* by IAB5 was found to be dependent on a 255-bp element in the proximal promoter [35]. Such proximal promoter elements that confer enhancer specificity have been described for several other genes [36]. Possibly related to the role of core-promoter elements in determining enhancer–promoter biochemical compatibility, it is becoming increasingly clear that alternative core promoter recognition complexes exist that can target enhancers and promoters in a tissue specific manner (for review see [37]). It is, therefore, an interesting possibility that different types of core promoter recognition complexes play a role in determining enhancer–promoter biochemical compatibility.

A recently developed genome-wide screening strategy identified thousands of functional enhancers that could activate an artificial promoter carrying both a TATA-box and DPE [38]. Remarkably, enhancers near ribosomal protein genes were under-represented in the identified set, which may be explained by the fact that ribosomal protein gene promoters carry a distinct initiator element (named TCT) [39] absent in the artificial promoter. It is, therefore, tempting to speculate that enhancers near ribosomal protein genes are compatible with TCT-containing promoters, but not with other promoters. In addition, it would be insightful to assess differential enhancer efficacies with this assay using promoters containing either a DPE or a TATA-box.

Although enhancer–promoter compatibility has not been as systematically studied in mammals, several examples suggest similar principles. The locus control region (LCR; a special, powerful enhancer) of the β -globin gene cluster interacts with only a single β -globin gene at a time, depending on the developmental stage of the embryo [40]. During one of these stages, the factors GATA-1 and FOG-1 are necessary to keep the LCR and the beta-major promoter together [41]. GATA-1 acts by recruiting its co-factor Lbd1, and tethering experiments have demonstrated that Lbd1 is responsible for bringing the LCR and promoter together [42]. Interestingly, when the LCR is inserted ectopically on a different chromosome, it preferentially contacts genes regulated by similar TFs, suggesting a degree of biochemical specificity; however, the parallel observation that ectopic LCR preferentially contacts active genes may confound this interpretation [43]. Also, during mouse erythroid cell differentiation, GATA-1 is directly responsible for switching of the *Kit* promoter from one enhancer to another approximately 150 kb away [44]. Finally, the activation of a minimal immunoglobulin gene promoter by its enhancer was shown to be dependent on promoter binding by a POU-domain-containing TF, providing insight into how this enhancer–promoter interaction is encoded [45].

In summary, these data illustrate that biochemical compatibility, often dictated by the presence or absence of a single protein, can be an important determinant of the selectivity of enhancer–promoter interactions.

Role of chromosome architecture

Naturally, physical contact between a distal enhancer and a promoter occurs only if the overall folding of the chromosomal fiber allows it. Interphase chromosomes are in constant Brownian motion [46–48] and can adopt many configurations in a partially random manner [49]. Therefore, many enhancer–promoter contacts may be inherently stochastic, in which case we should consider them in terms of contact frequencies.

Polymer physics theory and computer simulations suggest that the contact frequency between two loci generally decays with their distance along the linear chromosome, in a manner whose details depend on the assumed polymer model [50,51]. The shape of the decay function can be empirically estimated from 5C and Hi-C data [52,53]. Although it varies between cell types and species, the contact frequency between two typical human genomic loci is roughly inversely proportional with distance along the genome, such that contact frequencies at distances of 100 kb are only ~2% of those at 2 kb (Figure 2). Consistent with the idea of contacts being established more effectively over short distances, for thousands of *Drosophila* enhancers identified through a powerful multiplexed ectopic reporter assay [38,54], the expression of target genes in the natural genome context was found to correlate inversely with distance from these enhancers [54]. Similar behavior has been observed for enhancers bound by the TF NF- κ B [17] and for functional enhancer–promoter pairs inferred from eRNA and mRNA expression across different cell types [28]. Finally, in a recent study where almost 500 enhancers were matched to their target gene based on the similarity of their expression pattern in *Drosophila* embryos, 88% of enhancers targeted the first gene upstream, downstream, or the host gene [55].

Alternatively, there are many examples of functional enhancer–promoter interactions that act over long linear distances [56]. Such long-range interactions may require assistance from architectural proteins. Indeed, evidence is mounting that certain proteins can tether linearly distant loci together in 3D space. A prime candidate for this function is CTCF, a DNA-binding protein that recognizes a well-defined sequence motif. In erythroid progenitor cells, CTCF was shown to play a prominent role in bringing the beta-globin LCR closer to its target promoter by binding to several sites in the locus and forming multiple loops [57]. In mammalian cells, CTCF can interact with cohesin [58], a ring-shaped protein complex that has the potential to encircle multiple DNA strands. The combination of CTCF and cohesin is preferentially found at genomic locations that exhibit constitutive (cell-type invariant) looping contacts detectable by 5C. Knockdown of either CTCF or a cohesin subunit caused microscopically detectable disruption of such a loop [59]. The transcriptional mediator complex [59,60] and non-coding RNA (ncRNA) (e.g. [61–63]; for review see [64]) have also recently been implicated in the formation of looping contacts at large distances and, therefore, might also be determinants of specific enhancer–promoter interactions.

Topologically associated domains (TADs) represent another, albeit related, architectural feature that can help direct enhancers to the right target promoters. TADs are ~1 Mb chromosomal domains characterized by intra-domain contacts between many different sequences within the domain [65–68]. They are found throughout the mammalian genome. By definition, contacts between TADs are relatively rare [65,67]. Because 5C and Hi-C data

are typically sampled from millions of cells, it is not known how many encounters occur within a single TAD at any time, although computer simulations backed up by fluorescent *in situ* hybridization (FISH) data suggest that it may be only a few [69]. Nevertheless, TADs are thought to be relatively compact units of chromatin, as suggested by FISH experiments showing that the average spatial distance between two loci within the same TAD is smaller than between two loci in different TADs [67]. Comparisons of 5C and Hi-C data from different cell types suggest that TAD organization across the mammalian genome is largely unaltered during differentiation [65,67].

TAD architecture creates units of substantially reduced ‘search space’ within which enhancers and promoters may find each other more easily. While contacts between promoters and enhancers occur preferentially within a TAD [17,65], this could be due to circular reasoning because TADs are defined based on their preferential intra-domain contacts. It is possible that enhancer–promoter interactions collectively both drive and profit from TAD formation at the same time. Interestingly, removal of a boundary region between two TADs has been observed to result in the formation of one larger TAD [67]. It should be worthwhile to investigate whether, in this setting, enhancer–promoter contacts occur across the two former TADs, and whether this can lead to altered gene activity.

Insulator elements

Insulator elements were originally identified as DNA elements that can prevent the activation of a promoter by an enhancer when placed between them [70,71]. Some insulator elements can control local chromatin structure (e.g., by preventing spreading of heterochromatin), but that is beyond the scope of this review. Insulator elements are bound by specific DNA-binding factors. At least five distinct insulator complexes have been described in *Drosophila* [72,73]. In vertebrates, insulator function has primarily been attributed to CTCF-binding factor (CTCF) [59]. For comprehensive reviews, see [74,75]; in this review, we briefly highlight some essential features.

Currently, the predominant models for how insulators work are the looping model and the decoy model [74]. The looping model posits that two or more insulator sites physically interact with each other, thereby establishing loops that alter the 3D conformation of the chromatin fiber in a manner that affects the ability of enhancers to interact with promoters. Indeed, 5C and HiC data reveal frequent encounters between CTCF-bound regions [17,18,76]. In addition, a global ChIA-PET assay of CTCF in mouse ES cells revealed enhancer regions prevalent among looping interactions of CTCF-bound loci [76]. Importantly, depending on the precise context, such contacts can either block or promote enhancer–promoter contacts. In the decoy model, insulators interact directly with an enhancer or promoter, thereby interfering with enhancer–promoter encounters. For example, a recent 3C study found that the well-characterized *Drosophila* insulator *Gypsy* was able to insulate the *white* gene from the eye enhancer in part by directly interacting with it [77].

The IAB5 enhancer, discussed above, illustrates how insulators can provide an additional layer of regulation on top of biochemical compatibility to determine promoter–enhancer interactions. IAB5 has an intrinsic preference for TATA-containing promoters, but was

shown to prefer a DPE-containing promoter when separated from TATA-containing promoters by an insulator [34]. Paradoxically, at its endogenous locus, the IAB5 enhancer activates *Abd-B* across two insulator sites (Fab7 and Fab8) [35]. There are many examples of enhancer-blocking activity of insulators [73,74], but most of these experiments have been performed in artificial ectopic settings, leaving the question of how relevant the identified effects are for endogenous function. Moreover, recent global analyses have made it clear that many physical encounters in the genome occur across insulator sites, suggesting that the extent to which insulators block promoter–enhancer interactions is at least highly context-dependent (e.g. [18,29,65]).

Local chromatin composition

What would happen if a reporter gene driven by a particular enhancer–promoter pair, functional in its endogenous genomic context, were to be randomly inserted elsewhere in the genome? The local chromatin composition would vary with the site of integration, and as outlined above, the presence or absence of a single sequence-specific protein can alter the compatibility of an enhancer–promoter pair. Therefore, we propose that local chromatin composition may modulate enhancer–promoter interaction, both physically and functionally. In any case, it is clear that local chromatin composition can affect enhancer and promoter activity. Tissue-specific enhancers are generally only active when bound by *p300* [78], while H3K27ac distinguishes active enhancers from inactive or poised enhancers containing H3K4me1 alone [79,80]. Poised enhancers are associated with the repressive marks H3K9me3 and H3K27me3 [80,81]. Targeting of a histone demethylase to selected enhancers using TAL effectors in several instances led to specific downregulation of proximal genes [82], demonstrating the importance of the chromatin composition of enhancers.

On a scale beyond that of individual genes, approximately one-third of the genome is packaged into large chromatin domains (~100 kb up to several Mb) that interact with the nuclear lamina [83,84]. These lamina-associated domains (LADs) overlap at least partially with a subset of TADs [65,67]. Most endogenous genes embedded in LADs show no detectable transcription [83–85] and a highly multiplexed reporter assay showed that promoters ectopically inserted into LADs are approximately sixfold less active than when integrated in inter-LAD regions [86]. Because this apparent chromatin effect on the ectopic promoters does not fully explain the complete lack of transcription exhibited by most endogenous genes in LADs, an interesting possibility is that the chromatin composition or spatial architecture of LADs further suppresses enhancer–promoter contacts. In general, the effect of chromatin context on enhancer–promoter interplay is still a largely unexplored area of research.

Concluding remarks

It seems likely that the specificity of enhancer–promoter interactions is determined through a combination of biochemical compatibility, 3D chromosome organization, insulators, and perhaps local chromatin composition. An exciting road lies ahead towards a fuller understanding of the rules that underlie the complex network of enhancer–promoter

interactions, with the ultimate goal of being able to predict normal and perturbed enhancer–promoter interactions from genomic sequence alone.

To achieve a predictive understanding of enhancer–promoter compatibility, large-scale functional studies will be essential. Recently developed high-throughput reporter assays will make it possible to test thousands or even millions of enhancer–promoter combinations in carefully controlled contexts (e.g. [38,87]). Alternatively, the consequences of thousands of sequence variations in selected promoters and enhancers can be studied (e.g., [88–90]). Such systematic approaches combined with computational analyses will lead the way towards a better understanding of biochemical compatibility rules.

Further improvements of 3C-based method may offer ways to generate comprehensive maps of physical enhancer–promoter encounters at increased coverage, precision, and resolution. However, it is still a matter of debate whether 3C-based methods detect only close (nanometer-range) physical contacts of DNA elements, or whether these data also contain significant contributions from the formation of large (>100 nm) nucleoprotein aggregates generated by formaldehyde fixation [91,92]. To achieve full mechanistic understanding, we will need to know much more about the dynamics of the spatial encounters. At any given time, do they occur in the majority, or in a small fraction of the cells? How stable are the underlying interactions in an individual cell? For example, 3D FISH revealed that *Shh* expression in the posterior mesenchyme of the mouse limb bud correlates with *Shh* contacting (at light microscopy resolution) its enhancer. Interestingly, the contacts are found only in a subset of the cells in which *Shh* is expressed, an observation that might explain the occurrence of transient pulses of transcription [93]. Regardless, Hi-C and other assays for mapping aspects of nuclear architecture [83] will likely help us to understand how enhancer–promoter interactions are guided by the spatial organization of chromosomes.

Scalable approaches in which functional enhancer–promoter interactions are perturbed in a targeted manner are now within reach, and will help firmly establish causal functional relationships. Genome editing methods, particularly the CRISPR/Cas9 technology [94], provide an opportunity to systematically delete or inactivate substantial numbers of enhancers and thus functionally identify their target promoters in a native context. It will be illuminating to compare such data with predictions based on physical contacts and correlative analyses such as the one described above.

To investigate the role of chromatin context as a determinant of enhancer–promoter interaction, enhancer–promoter pairs could be functionally examined at thousands of random genomic integration sites [86], or thousands of enhancer–promoter pairs could be assayed by targeted integration into several distinct chromatin contexts [87]. In a complementary approach, the chromatin environment of endogenous enhancers or promoters could be altered through the targeting of specific histone-modifying enzymes [82]. These and other functional genomics approaches will pave the way towards a deeper understanding of the determinants of enhancer–promoter interactions.

References

1. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011; 144:327–339. [PubMed: 21295696]
2. Smith E, Shilatifard A. Enhancer biology and enhanceropathies. *Nat. Struct. Mol. Biol.* 2014; 21:210–219. [PubMed: 24599251]
3. Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 2012; 13:613–626. [PubMed: 22868264]
4. Lenhard B, et al. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* 2012; 13:233–245. [PubMed: 22392219]
5. Fromm G, et al. An embryonic stage-specific enhancer within the murine beta-globin locus mediates domain-wide histone hyperacetylation. *Blood*. 2011; 117:5207–5214. [PubMed: 21321362]
6. Cajiao I, et al. Bystander gene activation by a locus control region. *EMBO J.* 2004; 23:3854–3863. [PubMed: 15359275]
7. Gomez-Skarmeta JL, et al. *Cis*-regulation of achaete and scute: shared enhancer-like elements drive their coexpression in proneural clusters of the imaginal discs. *Genes Dev.* 1995; 9:1869–1882. [PubMed: 7649474]
8. Grossniklaus U, et al. The *Drosophila* sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes Dev.* 1992; 6:1030–1051. [PubMed: 1317319]
9. Nickol JM, Felsenfeld G. Bidirectional control of the chicken beta- and epsilon-globin genes by a shared enhancer. *Proc. Natl. Acad. Sci. U.S.A.* 1988; 85:2548–2552. [PubMed: 3357880]
10. Spitz F, et al. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell*. 2003; 113:405–417. [PubMed: 12732147]
11. Lettice LA, et al. Disruption of a long-range *cis*-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. U.S.A.* 2002; 99:7548–7553. [PubMed: 12032320]
12. Lower KM, et al. Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:21771–21776. [PubMed: 19959666]
13. Lomvardas S, et al. Interchromosomal interactions and olfactory receptor choice. *Cell*. 2006; 126:403–413. [PubMed: 16873069]
14. Dekker J, et al. Capturing chromosome conformation. *Science*. 2002; 295:1306–1311. [PubMed: 11847345]
15. van Steensel B, Dekker J. Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol.* 2010; 28:1089–1095. [PubMed: 20944601]
16. Zhang J, et al. ChIA-PET analysis of transcriptional chromatin interactions. *Methods*. 2012; 58:289–299. [PubMed: 22926262]
17. Jin F, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013; 503:290–294. [PubMed: 24141950]
18. Sanyal A, et al. The long-range interaction landscape of gene promoters. *Nature*. 2012; 489:109–113. [PubMed: 22955621]
19. Li G, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*. 2012; 148:84–98. [PubMed: 22265404]
20. Agelopoulos M, et al. Developmental regulation of chromatin conformation by Hox proteins in *Drosophila*. *Cell Rep.* 2012; 1:350–359. [PubMed: 22523743]
21. Andrey G, et al. A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science*. 2013; 340:1234167. [PubMed: 23744951]
22. Ghavi-Helm Y, et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*. 2014
23. Filion GJ, et al. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell*. 2010; 143:212–224. [PubMed: 20888037]
24. Moorman C, et al. Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 2006; 103:12027–12032. [PubMed: 16880385]

25. Gao F, et al. Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinformatics*. 2004; 5:31. [PubMed: 15113405]
26. Edelman LB, Fraser P. Transcription factories: genetic programming in three dimensions. *Curr. Opin. Genet. Dev.* 2012; 22:110–114. [PubMed: 22365496]
27. Thurman RE, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012; 489:75–82. [PubMed: 22955617]
28. Andersson R, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014; 507:455–461. [PubMed: 24670763]
29. Shen Y, et al. A map of the *cis*-regulatory sequences in the mouse genome. *Nature*. 2012; 488:116–120. [PubMed: 22763441]
30. Li X, Noll M. Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the *Drosophila* embryo. *EMBO J.* 1994; 13:400–406. [PubMed: 8313885]
31. Merli C, et al. Promoter specificity mediates the independent regulation of neighboring genes. *Genes Dev.* 1996; 10:1260–1270. [PubMed: 8675012]
32. Juven-Gershon T, et al. Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. *Genes Dev.* 2008; 22:2823–2830. [PubMed: 18923080]
33. Butler JE, Kadonaga JT. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev.* 2001; 15:2515–2519. [PubMed: 11581157]
34. Ohtsuki S, et al. Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes Dev.* 1998; 12:547–556. [PubMed: 9472023]
35. Akbari OS, et al. A novel promoter-tethering element regulates enhancer-driven gene expression at the bithorax complex in the *Drosophila* embryo. *Development*. 2008; 135:123–131. [PubMed: 18045839]
36. Kwon D, et al. Enhancer-promoter communication at the *Drosophila* engrailed locus. *Development*. 2009; 136:3067–3075. [PubMed: 19675130]
37. Goodrich JA, Tjian R. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev. Genet.* 2010; 11:549–558. [PubMed: 20628347]
38. Arnold CD, et al. Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science*. 2013; 339:1074–1077. [PubMed: 23328393]
39. Parry TJ, et al. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev.* 2010; 24:2013–2018. [PubMed: 20801935]
40. Tolhuis B, et al. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell*. 2002; 10:1453–1465. [PubMed: 12504019]
41. Vakoc CR, et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol. Cell*. 2005; 17:453–462. [PubMed: 15694345]
42. Deng W, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*. 2012; 149:1233–1244. [PubMed: 22682246]
43. Noordermeer D, et al. Variegated gene expression caused by cell-specific long-range DNA interactions. *Nat. Cell Biol.* 2011; 13:944–951. [PubMed: 21706023]
44. Jing H, et al. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol. Cell*. 2008; 29:232–242. [PubMed: 18243117]
45. Bertolino E, Singh H. POU/TBP cooperativity: a mechanism for enhancer action from a distance. *Mol. Cell*. 2002; 10:397–407. [PubMed: 12191484]
46. Chubb JR, et al. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr. Biol.* 2002; 12:439–445. [PubMed: 11909528]
47. Marshall WF, et al. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr. Biol.* 1997; 7:930–939. [PubMed: 9382846]
48. Nozaki T, et al. Flexible and dynamic nucleosome fiber in living mammalian cells. *Nucleus*. 2013; 4:349–356. [PubMed: 23945462]
49. Nagano T, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 2013; 502:59–64. [PubMed: 24067610]

50. Fudenberg G, Mirny LA. Higher-order chromatin structure: bridging physics and biology. *Curr. Opin. Genet. Dev.* 2012; 22:115–124. [PubMed: 22360992]
51. Tark-Dame M, et al. Chromatin folding – from biology to polymer models and back. *J. Cell Sci.* 2011; 124:839–845. [PubMed: 21378305]
52. Dekker J, et al. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* 2013; 14:390–403. [PubMed: 23657480]
53. Halverson JD, et al. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Rep. Prog. Phys.* 2014; 77:022601. [PubMed: 24472896]
54. Shlyueva D, et al. Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. *Mol. Cell.* 2014; 54:180–192. [PubMed: 24685159]
55. Kvon EZ, et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature.* 2014
56. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am. J. Hum. Genet.* 2005; 76:8–32. [PubMed: 15549674]
57. Splinter E, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 2006; 20:2349–2354. [PubMed: 16951251]
58. Merkenschlager M, Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell.* 2013; 152:1285–1297. [PubMed: 23498937]
59. Phillips-Cremins JE, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell.* 2013; 153:1281–1295. [PubMed: 23706625]
60. Kagey MH, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature.* 2010; 467:430–435. [PubMed: 20720539]
61. Lai F, et al. Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature.* 2013; 494:497–501. [PubMed: 23417068]
62. Li W, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature.* 2013; 498:516–520. [PubMed: 23728302]
63. Melo CA, et al. eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell.* 2013; 49:524–535. [PubMed: 23273978]
64. Lam MT, et al. Enhancer RNAs and regulated transcriptional programs. *Trends Biochem. Sci.* 2014; 39:170–182. [PubMed: 24674738]
65. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012; 485:376–380. [PubMed: 22495300]
66. Hou C, et al. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell.* 2012; 48:471–484. [PubMed: 23041285]
67. Nora EP, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature.* 2012; 485:381–385. [PubMed: 22495304]
68. Sexton T, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell.* 2012; 148:458–472. [PubMed: 22265598]
69. Giorgetti L, et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell.* 2014; 157:950–963. [PubMed: 24813616]
70. Kellum R, Schedl P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell.* 1991; 64:941–950. [PubMed: 1848159]
71. Kellum R, Schedl P. A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell. Biol.* 1992; 12:2424–2431. [PubMed: 1569958]
72. Gurudatta BV, Corces VG. Chromatin insulators: lessons from the fly. *Brief. Funct. Genomic Proteomic.* 2009; 8:276–282. [PubMed: 19752045]
73. Kyrchanova O, Georgiev P. Chromatin insulators and long-distance interactions in *Drosophila*. *FEBS Lett.* 2014; 588:8–14. [PubMed: 24211836]
74. Herold M, et al. CTCF: insights into insulator function during development. *Development.* 2012; 139:1045–1057. [PubMed: 22354838]
75. Phillips-Cremins JE, Corces VG. Chromatin insulators: linking genome organization to cellular function. *Mol. Cell.* 2013; 50:461–474. [PubMed: 23706817]

76. Handoko L, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 2011; 43:630–638. [PubMed: 21685913]
77. Kyrchanova O, et al. Effective blocking of the white enhancer requires cooperation between two main mechanisms suggested for the insulator function. *PLoS Genet.* 2013; 9:e1003606. [PubMed: 23861668]
78. Visel A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature.* 2009; 457:854–858. [PubMed: 19212405]
79. Creighton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* 2010; 107:21931–21936. [PubMed: 21106759]
80. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature.* 2011; 470:279–283. [PubMed: 21160473]
81. Zentner GE, et al. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* 2011; 21:1273–1283. [PubMed: 21632746]
82. Mendenhall EM, et al. Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.* 2013; 31:1133–1136. [PubMed: 24013198]
83. Guellen L, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature.* 2008; 453:948–951. [PubMed: 18463634]
84. Peric-Hupkes D, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell.* 2010; 38:603–613. [PubMed: 20513434]
85. Kind J, et al. Single-cell dynamics of genome-nuclear lamina interactions. *Cell.* 2013; 153:178–192. [PubMed: 23523135]
86. Akhtar W, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell.* 2013; 154:914–927. [PubMed: 23953119]
87. Dickel DE, et al. Function-based identification of mammalian enhancers using site-specific integration. *Nat. Methods.* 2014; 11:566–571. [PubMed: 24658141]
88. Kwasniewski JC, et al. Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl. Acad. Sci. U.S.A.* 2012; 109:19498–19503. [PubMed: 23129659]
89. Melnikov A, et al. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* 2012; 30:271–277. [PubMed: 22371084]
90. Patwardhan RP, et al. Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.* 2012; 30:265–270. [PubMed: 22371081]
91. Belmont AS. Large-scale chromatin organization: the good, the surprising, and the still perplexing. *Curr. Opin. Cell Biol.* 2014; 26:69–78. [PubMed: 24529248]
92. Gavrillov AA, et al. Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. *Nucleic Acids Res.* 2013; 41:3563–3575. [PubMed: 23396278]
93. Amano T, et al. Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell.* 2009; 16:47–57. [PubMed: 19097946]
94. Mali P, et al. Cas9 as a versatile tool for engineering biology. *Nat. Methods.* 2013; 10:957–963. [PubMed: 24076990]
95. Palstra RJ. Close encounters of the 3C kind: long-range chromatin interactions and transcriptional regulation. *Brief. Funct. Genomic Proteomic.* 2009; 8:297–309. [PubMed: 19535505]

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.

Highlights

- Enhancers can interact with distal promoters in a highly selective manner.
- High-throughput technologies enable global mapping of these interactions.
- Selectivity mechanisms range from biochemical compatibility to 3D chromatin structure.
- We discuss future experiments towards a more complete understanding of these issues.

Figure 1.

Distinct mechanisms can drive promoter–enhancer interaction specificity. (A) Biochemical compatibility. Enhancer E2, but not E1, is compatible with the target promoter (P), explaining the selective interaction of E2 with P. (B) Spatial architecture. While E2 and E3 are both biochemically compatible with the promoter, spatial architecture dictates that only E2 interacts with the promoter. (C) Insulation. Although E3 is compatible with the promoter, their interaction is blocked by the insulator (Ins), possibly by alteration of the 3D structure or because the insulator serves as a decoy (not shown; see main text). (D) Chromatin environment. P selectively contacts E1 because the compatibility of E2 with P has been altered by its chromatin environment. This latter mechanism is still speculative.

Figure 2.

The frequency at which two loci encounter each other in 3D space, as estimated from 5C data, is inversely proportional to their linear distance along the genome. The average was calculated based on all loci assayed in [18], and separately for each of the three cell-types studied. Since no filter for statistical significance at the level of individual locus–locus pairs was applied, this analysis represents the average trend of all interactions, including nonspecific ones [18]. Note how contact frequencies at distances of 100 kb are only ~2% of those at 2 kb (dotted line), except for in HeLa cells (~10%).