Close encounters of the 3C kind: long-range chromatin interactions and transcriptional regulation

Robert-Jan Theo Sijtse Palstra

Advance Access publication date 27 July 2009

Abstract

The transcriptional output of genes in higher eukaryotes is frequently modulated by cis-regulatory DNA elements like enhancers. On the linear chromatin template these elements can be located hundreds of kilobases away from their target gene and for a long time it was a mystery how these elements communicate. For example, in the β-globin locus the main regulatory element, the Locus Control Region (LCR), is located up to 40–60 kb away from the β-globin genes. Recently it was demonstrated that the LCR resides in close proximity to the active β-globin genes while the intervening inactive chromatin loops out. Thus the chromatin fibre of the β-globin locus adopts an erythroid-specific spatial organization referred to as the Active Chromatin Hub (ACH). This observation for the first time demonstrated a role for chromatin folding in transcriptional regulation. Since this first observation in the β-globin locus, similar chromatin interactions between regulatory elements in several other gene loci have been observed. Chromatin loops also appear to be formed between promoters and 3'UTRs of genes and even trans-interactions between loci on different chromosomes have been reported. Although the occurrence of long-range chromatin contacts between regulatory elements is now firmly established it is still not clear how these long-range contacts are set up and how the transcriptional output of genes is modified by the proximity of cis-regulatory DNA elements. In this review I will discuss the relevance of interactions between cis-regulatory DNA elements in relation to transcription while using the β-globin locus as a guideline.

Keywords: long-range activation; chromosome conformation capture; β-globin; transcription

LONG-RANGE ACTIVATION OF GENE TRANSCRIPTION: AN INTRODUCTION

Transcription in eukaryotes is initiated at DNA elements called promoters. The promoter core consists of a varying combination of elements like e.g. the TATA box, initiator element and downstream core promoter element. The core promoter positions and orientates the pre-initiation complex (PIC), containing the RNA polymerase II in complex with a variety of general transcription factors, and ensures the exact initiation and direction of the RNA transcript. The promoter is complemented by upstream-binding sites for (tissue-specific) transcription factors which make the promoter accessible and stimulate the assembly of the PIC [1]. In general, the promoter contains all the information necessary for the correct temporal and spatial expression of the linked gene however gene expression can be boosted to very high levels by the presence of cis-regulatory elements called enhancers. Enhancers are DNA sequences containing binding sites for (ubiquitous and tissue-specific) transcription factors which are located far (up to 100 kb) upstream or downstream relative to the gene they regulate. How these enhancers communicate with the promoters of the regulated gene has long been an issue of intense debate with two models prevailing: the non-contact and contact models. In the non-contact model a ‘signal’, often envisioned as a protein complex or...
RNA polymerase II, initiates at the enhancer and travels along the chromatin fiber until it reaches the to be activated promoter. In the contact model the enhancer, with bound proteins, travels through the nuclear space and directly contacts the promoter whereby the intervening chromatin loops out. The development of new molecular techniques like 3C (Chromatin Conformation Capture) and RNA TRAP (RNA Tagging and Recovery of Associated Proteins) and their application to the mouse β-globin locus first demonstrated that β-globin enhancers indeed reside in close proximity to the active β-globin genes [2–4]. This observation has since been confirmed in many other loci suggesting that intra-chromosomal interactions between chromosomal elements are a common theme. Recent adaptations of 3C technology to allow genome wide detection of chromatin interactions have even revealed several intra- and inter-chromosomal interactions between genomic loci, although their functional relevance is currently unclear and remains controversial. The inter-chromosomal interactions have recently been the subject of several excellent reviews and will not be further discussed in this review [5–8].

LONG-RANGE INTERACTIONS IN THE β-GLOBIN LOCUS: A BRIEF HISTORY

The β-globin locus has a long history as the model system of choice to study biochemical, genetic and molecular biological processes and many ground breaking discoveries have been made using this biological system [9]. The mouse and human β-globin loci are similarly organized and consist of several (four to five) developmentally regulated erythroid-specific genes. The genes are organized in a sequential, linear fashion along the chromatin fiber roughly reflecting the order of their expression during ontogeny. The produced β-globin protein together with α-globin protein (expressed from a different locus on a different chromosome) forms a stage specific hemoglobin complex which is tailored to the oxygen tension prevalent at the relevant developmental stage. The β-globin genes are expressed at extremely high levels and as a consequence β-globin, as part of hemoglobin, is the most abundant protein in erythroid cells. This high level expression is achieved by the presence of a very strong enhancer element, the locus control region (LCR) upstream of the genes. The presence of this element was first inferred in β-thalassemia patients that had all β-globin genes intact but possessed a deletion of a region upstream of the genes [10]. Consequent studies identified several erythroid specific DNase I-hypersensitive sites in this region and inclusion of these hypersensitive sites in a transgenic β-globin construct resulted in high-level β-globin expression, independent of the site of integration, i.e. refractory to silencing, in transgenic mice [11]. Apart from the DNase I-hypersensitive sites of the LCR other DNase I-hypersensitive sites are found both upstream as downstream of the genes [12, 13].

How the LCR is able to activate the β-globin genes over a large distance, up to 40–60 kb, was the subject of intense study in the next 15 years. Although the outcome of several elegant experiments like e.g. the observations that only one β-globin gene is upregulated at a time by the LCR and that β-globin genes compete for activation by the LCR, were most easily explained by a looping model of enhancer action [14–17], the matter remained controversial [18].

Since the size of the β-globin locus is below the resolution of current light microscopy direct visualization of loop formation in the β-globin locus was not possible. Fortunately the simultaneous development of two independent biochemical approaches, RNA TRAP and 3C, provided a means to map the spatial organization of gene loci [2–4]. RNA TRAP is a clever adaptation of the RNA FISH protocol in which is in close proximity to the nascent RNA transcript is tagged with biotin, which is after fragmentation purified from untagged chromatin using streptavidin and consequently detected using PCR. When applied to the mouse β-globin locus not only the active β-globin gene was recovered but also sequences in the LCR were readily detected while intervening sequences were less abundant [3]. These results demonstrated that the LCR resides in close proximity to the active gene while the intervening chromatin loops out.

3C was originally developed to map Yeast chromosomes [2]. In 3C chromatin is formaldehyde crosslinked followed by fragmentation using a restriction enzyme. The chromatin fragments are re-ligated under conditions that favor intra molecular ligations, i.e. chromatin fragments that were in spatial close proximity to each other and therefore are cross-linked to each other are preferentially ligated. After de-crosslinking and purification the ligated DNA
fragments are detected by PCR. This method was adapted to the mouse β-globin locus and high cross-linking frequencies between a restriction fragment containing the β-major promoter and DNA fragments in the LCR were detected while intervening restriction fragments displayed lower crosslinking frequencies [4]. These observations were very similar to the results that were obtained in the RNA TRAP study. 3C however, unlike RNA TRAP can be applied to non-transcribed loci or elements and crucially, when the β-globin locus was analyzed in fetal brain were this locus is not active no interaction between the β-major gene and the LCR was detected. This observation unequivocally demonstrated that enhancer-gene interactions correlate with the transcriptional activity of the gene.

Analysis of the chromatin interactions of other genetic elements within the β-globin locus demonstrated that the active β-globin locus folds into an intricate erythroid specific conformation called an active chromatin hub (ACH) [4]. In a subsequent study the folding of the β-globin locus in erythroid progenitors and erythroid cells at a different developmental stage which express different β-globin genes was investigated [19]. This study revealed that in erythroid progenitors a 'scaffold' structure is formed between a subset of distal DNase I-hypersensitive sites which was referred to as a chromatin hub (CH) and that later during differentiation LCR–gene contacts are made with the β-globin gene which is active at that specific stage. Moreover the authors were able to demonstrate that the human β-globin locus, as a transgene in mice, folds in a similar fashion as the mouse β-globin locus [19].

**CHROMATIN LOOPS: AN EXCEPTION OR A COMMON THEME?**

In recent years 3C, or adaptations thereof, has become the method of choice to study chromatin interactions in the nucleus and map the folding of complex gene loci in mammalian cells. Although 3C is in principle a straight forward and robust method one should realize though, that 3C measures steady-state average conformations across a population of cells and that the meaningful interpretation of 3C results depends critically on the proper controls [20, 21]. In recent years a steady trickle of studies, of diverse quality, using 3C to map a wide variety of gene loci has appeared in literature [22]. Combined these studies support the generality of chromatin loop formation in the regulation of nuclear processes although variations on the theme are evident and some interesting examples will be discussed next.

**Activating Chromatin Loops**

In the imprinted mouse Igf2/H19 locus the Igf2 gene which encodes insulin-like growth factor, is only expressed from the paternal allele. Conversely the non-coding H19-transcript is only expressed from the maternal allele. In this locus chromatin loops are established between the Igf2/H19 enhancer and either the Igf2 or H19 gene all depending on the methylation status of the DMR which is located between the Igf2/H19 enhancer and the Igf2-gene [23].

The mouse T112 locus contains three cytokine genes IL4, IL5, IL13. Interestingly the TH2 LCR is located within the ubiquitously expressed RAD50 gene. An elegant 3C study demonstrated that the promoters of the cytokine genes are in close proximity to each other in four different but closely related cell-types. Crucially in CD4+ T cells and natural killer cells, but not in B cells and fibroblasts the T112 LCR participates in these interactions [24]. Deletion of one hypersensitive site from the T112 LCR results in reduced expression of the cytokine genes and the interaction between the T112 LCR and the promoters of these genes is reduced [25]. The major activating element of the α-globin locus also resides in a housekeeping gene and an intricate pattern of interactions between regulatory elements has been detected in mouse and chicken cells [26–28].

**Dynamic Chromatin Loops**

As mentioned before 3C measures steady state average chromatin conformations in a population of cells. This raises the question how dynamic the measured interactions are. Gene activation by nuclear receptors like e.g. the estrogen receptor often proceeds in a cyclical fashion [29, 30]. In several instances looping between nuclear receptor response elements and activated promoters has been detected [31–35]. In the case of the vitamin D responsive p21 gene cyclical interactions between different vitamin D responsive elements (VDREs) have indeed been detected which correlates with peaks of phosphorylated RNA pol II occupancy and gene transcription [35]. Interesting after the first two cycles no further
interactions between VDREs and the promoter are detected although a clear third peak of phosphorylated RNA pol II occupancy and gene transcription is observed.

Chromatin Loops at Ectopic Locations
Placement of transcriptional enhancers at ectopic locations e.g. via a translocation event or viral integration of a gene therapy vector often results in the aberrant transcription of genes. This upregulation can take place via chromatin loop formation as was recently demonstrated for the t(14;18) lymphoma. In this translocation the IgH 3’ enhancer is placed at a 350-kb distance from the bcl-2 gene which becomes upregulated. 3C analysis revealed a physical interaction between the IgH 3’ enhancer and the bcl-2 promoter that was responsible for the increased bcl-2 expression in cells harboring this translocation [36]. A different example of long-range activation via chromatin looping of an enhancer at an ectopic location was provided by an experiment where the human β-globin LCR was targeted to a gene dense region on mouse chromosome 8. Several genes were upregulated in erythroid cells when the LCR was present and this correlated with increased looping between the site of integration and the upregulated genes [37].

Chromatin Loops between Promoters and Terminators
Chromatin loops are not only found between distal regulatory elements and promoters but also between promoters and the terminator regions of genes. Physical interaction between promoters and terminators was already suggested based on observations that the 3’ end processing machinery and initiation machinery are functionally and physically linked (e.g. [38, 39]). Interactions between promoters and terminator regions, thereby forming gene loops, were recently indeed detected in yeast and are thought to facilitate re-initiation of consecutive rounds of transcription [39–41]. Later these observations were extended to the mammalian mitochondrial rRNA genes, the integrated HIV pro-virus and the human CD68 gene [42–44]. A more complicated form of gene looping was found to be present in the un-induced BRCA1 gene were besides a transcription dependent promoter termination loop several interactions between the promoter and specific intronic and exonic regions were observed [45]. Interestingly upon induction of BRCA1 expression with estrogen (which involves cell proliferation) the promoter–terminator loop, but not the promoter–intronic/exonic loops, was lost.

Chromatin Loops in Repression
Chromatin looping not only plays a role in gene activation but also in gene inactivation as is evident from several recent reports. In the imprinted gene cluster carrying the Dlx5 and Dlx6 genes a repressive 11-kb loop is formed. In Mecp2-null mice this repressive loop is abrogated and replaced by an interaction between Dlx5 and far distal sequence resulting in activation of the gene [46]. The Kit locus is active in immature erythroid cells and an activating chromatin loop is formed between an upstream enhancer and the promoter. Upon maturation this loop is replaced by a loop with a down stream element resulting in repression of the gene [47].

Transcription programs which are mitotic heritable are transmitted by Polycomb group proteins bound to Polycomb response elements (PREs). The PREs and promoters of the Drosophila homeotic locus bithorax complex in its repressed state interact at a distance to form a higher order structure that mediates the epigenetic silencing of this locus [48]. Similarly, in human cells the GATA-4 locus displays a multi-loop organization that is mediated by Polycomb group proteins, H3K27me3 and DNA hypermethylation when the locus is in its repressed state. When GATA-4 is active, DNA is hypomethylated, tri-methylation of H3K27 is lost and binding of Polycomb group proteins is strongly reduced and as a consequence the multi-loop organization is resolved allowing transcription to commence [49].

Structural Chromatin Loops
Some of the chromatin loops that have been detected don’t have a clear function in transcriptional regulation. As mentioned before the β-globin locus contains several hypersensitive sites upstream of the LCR and downstream of the genes (in mouse these are HS-85, HS-60/62 and 3’HS1, respectively) and loops can be detected between these sites in erythroid progenitor as well as mature erythroid cells [19]. All these sites together with HS5 of the LCR contain binding sites for the transcription factor CTCF which mediates the loop formation. Deletion of CTCF results in a reduction of the chromatin loops formed between these sites [50]. Furthermore, deletion or mutation of these hypersensitive sites doesn’t have a noticeable effect on
adult stage β-globin gene expression [50–53]. Therefore it has been suggested that these sites have a structural function in organizing the chromatin fibre.

Chromatin Loops between Distal Gene Loci

When the transcription of several selected loci on mouse chromosome 7 (Eraf, Uros, Igf2, Kcnq1ot1) was assessed using 3D RNA-FISH it was noted that a high proportion of the signals colocalized with the β-globin signal. Subsequent DNA-FISH and immunofluorescence analysis demonstrated that the colocalization was dependent on transcription of the genes and colocalized to RNA polymerase II foci. The interaction of the β-globin locus with these distal genes was confirmed using a 3C assay. The authors suggested that widely separated genes are dynamically recruited to shared transcription factories [54]. Eraf, Uros and Kcnq1ot1 were also identified as interacting with the active β-globin locus in a study which used an adapted 3C approach to perform an unbiased search for DNA loci that interact with the β-globin locus. However this study also demonstrated that these interactions are not confined to just these genes but are more spread out and encompass chromatin regions of about 150 kb in size containing multiple active genes. Moreover when the β-globin locus is inactive it still displays long-range interactions with other genomic loci but this time these regions tend to be transcriptionally inactive [55]. Whether these chromatin loops between distal gene loci are governed by regulatory elements and have any functional relevance or are a mere consequence of chromosome folding remains to be established.

In conclusion, chromatin looping is involved in many chromatin related processes ranging from transcriptional activation, repression, enhancer blocking, recombination events, to structural organization of the genome. It is worth to mention that although looping is definitely involved in transcriptional activation and repression other mechanisms like spreading processes are likely to co-exist. The best known example is the spreading of repressive heterochromatin which is responsible for position effect variegation [56, 57]. Another example is the observation that the activation of the human ε-globin gene by the LCR seems to depend on a processive mechanism from the LCR to the ε-globin gene [58–60].

HOW TO KEEP IN TOUCH?
FACTORS INVOLVED IN MEDIATING LONG RANGE INTERACTIONS

How chromatin loops are formed is still unknown but several factors have been implicated in the maintenance of chromatin loops. Several transcription factors have the ability to multi-merize and indeed the transcription factors SP1 and Bach1/MafK are able to induce DNA loops between cognate binding sites in an in vitro system [61, 62]. Elegant experiments demonstrated that the Drosophila GAGA factor is able to bridge two distant DNA elements and in such a way triggers enhancer-dependent gene transcription [63]. Interestingly this GAGA-induced enhancer-dependent gene transcription also takes place when the promoter and enhancer reside on separate plasmids.

Soon after chromatin loops had been discovered in the β-globin locus investigations into which factors were responsible in maintaining these loops ensued. A large body of data generated over the years by many groups have implicated several proteins in β-globin activation [64, 65].

EKLF

The erythroid kruppel like factor (EKLF) was one of the first transcription factors whose role in chromatin looping was investigated. EKLF is a zinc-finger protein which binds to CACCC elements located in the β-major promoter and HS2 and HS3 of the LCR. EKLF is indispensable for adult β-globin expression and mice lacking EKLF die around E14 [66, 67]. The failure to express β-major is reflected in the fact that strongly reduced interactions between the LCR and promoter of β-major are observed [68]. Interactions between the distal DNase I-hypersensitivity sites remain intact and the β-globin locus in EKLF-null erythroid cells resembles the structure found in erythroid progenitor cells [19, 68]. LCR-β-major gene contacts could be restored in the EKLF-null mice by inducing a transgenic EKLF-fusion protein to relocate from the cytoplasm to the nucleus and importantly this took place even in the absence of protein synthesis demonstrating the direct involvement of EKLF in loop formation [68].

GATA-1

GATA-1 regulates red cell development and binds several high affinity GATA motifs present in
the β-globin locus [69]. Interactions between the β-globin LCR and the β-major promoter were found to be strongly impaired in a GATA-1 null erythroid cell-line and re-introduction of an activatable GATA-1 restored loop formation and β-globin transcription independent of protein synthesis [70]. Unfortunately, interactions between the distal DNase I-hypersensitive sites were not assayed in this study. Reintroduction of a mutated GATA-1 protein, which abolishes an interaction with its cofactor FOG-1, fails to restore the interaction between the LCR and β-major gene demonstrating a function for the FOG-1-GATA-1 interaction in loop formation [70]. Interestingly GATA-1 is also implicated in the formation of a repressive loop in the Kit locus. In this case a GATA-2 dependent activating loop in immature erythroid cells is replaced by a different GATA-1 dependent loop in mature erythrocytes which represses Kit expression. As is the case in the β-globin locus, formation of the GATA-1 dependent loop in the Kit locus requires the cofactor FOG-1 [47]. It is interesting to note that GATA-1 is able to form distinct repressive as well as active complexes in erythroid cells [71].

NF-E2

Another intensively studied transcription factor is the heterodimer NF-E2 although its precise role in β-globin activation remains controversial. NF-E2 consists of two DNA-binding subunits; a ubiquitous MafK protein and the erythroid specific p45 NF-E2 partner. Indications that p45 NF-E2 might be important for β-globin activation comes from a cell line which harbors a retroviral integration in the Fli-2 locus [72]. These cells do not express p45 NF-E2 anymore and as a consequence cannot sustain high level β-globin expression. Conversely p45 NF-E2 null mice are viable and only display mildly reduced levels of β-globin expression [73]. 3C analysis of these p45 NF-E2 null mice revealed no major reduction in interactions between the LCR and β-globin gene [74]. Increased binding of the p45 NF-E2 related factor Nrf2 at HS2 was observed which could potentially compensate for the lack of p45 NF-E2 although mice which lack both p45 NF-E2 and either Nrf2 or Nrf3 display no additional erythroid phenotype i.e. they express β-globin at near normal levels [75–77]. A recent study in a mouse erythroid leukaemia cell line approached this intriguing problem using a RNAi knockdown strategy against the MafK subunit [78]. Knockdown of MafK resulted in reduced MafK protein levels although still a substantial amount of MafK as well as p45 NF-E2 is detected at several regulatory elements. Nevertheless after induction of these MafK knockdown cells β-major expression is strongly reduced and this is reflected in a reduced interaction between HS2 and the β-major promoter [78].

Transcription factors involved in β-globin switching

During development the human β-globin locus switches its expression from fetal (γ-genes) to the adult genes (β- and δ-genes). This switching is delayed in Ikaros-null mice that carry the complete human β-globin locus as a transgene [79]. Ikaros strongly binds HS3 of the human β-globin LCR as well as regions just upstream of the γ-globin and δ-globin genes [80, 81]. Both groups detected a delayed switch in β-globin gene expression in the absence of functional Ikaros which is reflected in a change in loop formation, whereby a loop between the human β-globin LCR and adult stage β-globin gene is replaced by a loop between the LCR and the fetal γ-globin genes although the groups differ in their interpretation for the cause of this [80, 81]. Keys et al use an Ikaros mutant that lacks the capacity to bind DNA and they conclude that Ikaros is involved in facilitating a loop between the β-globin LCR and a region upstream of the δ-gene [80]. Bottardi and colleagues demonstrate a functional interaction between Ikaros and GATA-1 which results in the recruitment of a repressive complex to the γ-gene promoter region. The authors suggest that this repressive complex reduces the interaction between the LCR and γ-genes which consequently results in an increased LCR β-globin gene promoter interaction [81]. This model is in agreement with previous data that demonstrates autonomous silencing of the γ-globin genes. Interestingly in this regard is a recent study which replaced the Aγ-gene promoter in a transgenic human β-globin locus by the stage independent porphobilinogen promoter. This substitution leads to failure to autonomous silence γ-gene expression and results in increased loop formation between the LCR and the γ-genes at the expense of the LCR-β-globin interaction [82].

Transcription factors at the Th2 cytokine locus

The interactions between the promoters and LCR of the Th2 cytokine locus is mediated by the
transcription factor STAT6 although a core configuration remains in the absence of STAT6 [24]. Interestingly, artificial expression of another transcription factor GATA-3, in combination with calcium ionophore treatment, in fibroblast resulted in chromosomal interactions in the T_{h2} cytokine locus although no cytokine expression was detected [24]. SATB1, a protein that binds specific AT-rich sequences known as S/MARs, is also implicated in the complete folding of the T_{h2} cytokine locus [83]. Recently this protein was also shown to be involved in the formation of chromatin loops in the β-globin locus [84].

CTCF

As mentioned previously the β-globin locus contains several distal DNase I-hypersensitive sites which contain binding sites for CTCF. CTCF (CCCTC-binding factor) is well known for its function as a mammalian insulator protein which is able to block activation of a promoter by an enhancer when CTCF binds between these elements [85]. The distal DNase I-hypersensitive sites in the β-globin locus interact with each other [4, 19], and this depends on CTCF binding to these regions [50]. A similar role for CTCF in loop formation has been detected in the Igf2/H19 imprinted locus [86]. Interestingly when an ectopic CTCF site (in the form of β-globin HS5) is placed between the LCR and the downstream genes a CTCF dependent new loop between HS5 and the newly introduced HS5 is formed which topologically isolates the LCR and nullifies LCR function [87]. A strong correlation between CTCF binding and the localization of cohesin at several loci, which seems to be important for the insulator function of CTCF, has been observed [88–91]. Interestingly, it was recently demonstrated that chromatin loops in the human apolipoprotein gene cluster depend on CTCF as well as cohesion binding to insulator elements [92]. Moreover, recruitment of cohesin to CTCF sites can be regulated in a developmental stage specific manner [93], providing an additional way to establish stage-specific chromatin loops.

Multi-protein complexes

The above mentioned examples illustrate that several DNA-binding proteins are involved in maintaining chromatin loops. However, it is unlikely that a single DNA-binding factor can be identified that is solely responsible for setting up or maintaining a chromatin loop. Several DNA-binding proteins bind simultaneously and often interdependently in clusters at regulatory DNA-elements. Moreover, DNA-binding proteins are often part of or recruit large protein complexes [71, 94–96]. For example GATA-1 interacts with a protein complex containing the ubiquitous non-DNA-binding protein Ldb1 [71, 94]. Ldb1 is recruited to HSi-4 of the LCR and the β-major gene in erythroid cells and knockdown of Ldb1 in these cells results in a reduced interaction between HS2 of the LCR and the β-major gene [97]. Another example of protein complexes which are recruited by transcription factors are chromatin modifying and remodeling complexes. GATA-1 for example interacts with the histone acetyltransferase CBP/p300 and the chromatin remodeler BRG1 and both proteins are found on the β-globin promoter as well as on several hypersensitive sites of the LCR [69, 98, 99]. Using a hypomorphic allele of BRG1, Kim and colleagues recently demonstrated that BRG1 is involved in chromatin loop formation in the β-globin locus although it is unclear if BRG1 functions directly in this process or has a more supporting function e.g. facilitating the binding of factors to the chromatin [100]. CBP/p300 is required for long-range activation of the β-globin genes by the LCR although a role in loop formation has not yet been demonstrated [101]. Besides the histone acetyltransferase activity of CBP/p300 these proteins also act as a scaffold for building protein complexes connecting different sequence specific transcription factors [102], which could conceivably be bound to different widely spaced DNA-elements. Interestingly CBP/p300 is also bound by EKLF [103], and NF-E2 [104]. The Mediator complex is another large protein complex which is contacted by different transcription factors and RNA polymerase II [105] and is found at the enhancers of active yeast genes [106]. One of the Mediator subunits Med1/Trap220 interacts with GATA-1 [107], although it is still unclear if Mediator is recruited to the β-globin locus. A recent study demonstrated that EKLF is able to bind the general transcription factor TAF9 [108]. This observation makes it conceivable that direct interactions between distally bound transcription factors and the PIC mediate long-range chromatin interactions.

It has also been proposed that RNA polymerase II itself is responsible for establishing chromatin loops between enhancers and promoters (e.g. [109]). RNA polymerase II molecules bound to enhancers
(which are often transcribed) and those bound to the promoters of genes would aggregate in so called transcription factories. However inhibition of transcription in erythroid cells doesn’t result in the disruption of the chromatin loops between the LCR and the β-globin genes even though binding of RNA polymerase II to these elements was severely reduced [110, 111]. In the light of these data a direct role for RNA polymerase II in maintaining chromatin loops seems unlikely although a possible role in setting up these loops in first instance remains possible.

HOW TO GET FROM A TO B; TAKE THE TRAIN OR A PLANE?
Currently two models prevail to describe how enhancers find their target promoters: the facilitated looping model and the random collision model. In the facilitated looping model the enhancer tracks along the chromatin fibre to the promoter of a gene. The facilitated looping model is particularly attractive since it provides a satisfying explanation for how enhancer blockers work. Enhancer blockers impair the activation of a gene by an enhancer when placed between these elements but not when placed upstream or downstream of these elements. It is easy to envision that a tracking enhancer could be stopped in its tracks when it encounters an enhancer-blocking element and its associated proteins. The mechanism by which an enhancer would track along the chromatin fibre is less clear. The fact that intergenic transcripts are detected between the LCR and β-globin genes in the β-globin locus led to the suggestion that RNA polymerase II might be involved [112, 113]. Distinct intermediate interactions between the LCR and intergenic sequences haven’t been detected in the β-globin locus, possibly because detection of these interactions would probably require a carefully executed time course experiment in a synchronized pool of cells. In CaCo-2 cells however, intermediates that could reflect a tracking enhancer have been detected in the HNF-4α cluster [114]. Alternatively an enhancer can be envisioned to be actively guided through the 3D nuclear space to the target promoter via a nuclear Actin network. Recently several studies assigned a role for the nuclear Actin/Myosin-I machinery in the movement of chromatin loci through the nuclear space [31, 115, 116], although their role in enhancer-promoter interactions remains unclear.

In the random collision model the flexibility of the chromatin fibre allows the enhancer to sample a proportion of the nuclear space and will thereby randomly encounter the chromatin fibre. The transient chromatin loop is stabilized upon encountering a high affinity site like a promoter bound by the proper transcription factors. There is currently no satisfactory explanation for how an enhancer blocker could prevent the enhancer from approaching a promoter through the 3D nuclear space. The random collision model is however better able to explain several experimental observations which are difficult to explain by the facilitated tracking model. For example at a given time only one single gene of the β-globin locus is active and the β-globin genes compete for the LCR. Transcription of the β-globin genes takes place in an alternate, flip-flop fashion whereby a gene proximal to the LCR can be transcribed after a transcription round of the more distal gene [15–17]. Furthermore, transcription of a marked human β-globin gene is favored over the endogenous β-globin gene when placed proximal to the LCR. However this advantage is lost when the marked β-globin gene is placed more proximal to the endogenous gene [14]. Finally enhancers are able to activate genes even when placed on separate DNA strands as is evident from the phenomena of transvection in Drosophila [117] and numerous in vivo and in vitro experiments (e.g. [63, 118]).

WE ARE CLOSE, BUT WHAT TURNS US ON?
It is currently unclear by which mechanism the proximity of an enhancer to the promoter would increase the transcriptional output of a gene. Early studies on the human β-globin LCR suggested that its main function was opening up of chromatin domains for transcription [119]. Although the domain opening capacity is definitely a feature of LCRs, especially the human β-globin LCR, this is not necessarily related to its enhancer function and chromatin looping. For example when the mouse LCR is deleted, the β-globin genes are expressed albeit at a low level and certain active chromatin marks remain present in the locus [120, 121]. Similarly in erythroid progenitor cells the LCR is not yet in proximity to the β-globin genes these genes are expressed at a low level and several active chromatin marks are already present [74, 122]. Chromatin looping between an enhancer and
promoter therefore doesn’t act like an on/off switch but increases the transcriptional output of a gene. It has been suggested that clustering of regulatory DNA elements in an active chromatin hub would increase the local concentration of the associated transcription factors which would lead to more efficient transcription [4]. Other groups have suggested that LCRs/enhancers deliver RNA polymerase II and basal transcription factors to the promoter [28, 123]. Looping between LCRs/enhancers and promoters could possibly also deliver other factors like elongation factors or protein modification factors (e.g. kinases) to the transcription unit. Interestingly an elegant study using LCR knockout mice suggested that the LCR primarily functions in transcriptional elongation [124]. An attractive option is that the LCR-promoter interaction stabilizes the PIC, possibly by stabilizing promoter terminator interactions [40, 41] which would lead to more efficient rounds of re-initiation [39].

It has also been suggested that enhancers target the associated genes to transcriptional competent nuclear compartments, such as e.g. transcription factories [109, 112]. It was indeed demonstrated that the mouse β-globin locus relocates from the nuclear periphery to the nuclear interior upon activation and importantly this re-localization is dependent on the presence of an intact LCR [125]. However, transcription of the β-globin genes is not dependent on this re-localization to the interior of the nucleus since transcription of the β-globin locus can be detected at the periphery.

Clearly, not much is known about the exact mechanisms by which long-range interactions between enhancers and promoters modulate transcription and much more data is needed to progress from pure speculation into solid model building.

FUTURE PERSPECTIVES
New genome wide techniques prove to be very powerful in identifying distal acting regulatory DNA elements like enhancers or nuclear receptor response elements [32, 126–128]. Progress in these technologies, their application to different cell types and their combined use in projects like the ENCODE project are likely to generate a large library of distal acting regulatory DNA elements in the near future. It should be possible to map the interactions between these regulatory elements and their target genes using large scale application of 3C technology and its derivatives especially when combined with massive parallel sequencing [6]. More challenging however is elucidating how interactions between regulatory DNA elements and promoters are set up and by which mechanisms these interactions influence the transcriptional outcome of the target genes. New model systems that allow the manipulation of chromatin loops without affecting developmental programs and the recently developed ‘reverse-ChIP’ methods [129, 130], that allow the specific isolation of proteins associated with specific genomic loci will be instrumental in addressing these questions. The years following the first demonstration of in vivo chromatin loops have proven to be 7 years of plenty. Although challenging, the next 7 years are likely to provide us with many new exciting insights into the mysteries of long-range chromatin interactions.

Key Points
- The transcriptional output of genes in higher eukaryotes is frequently modulated by cis-regulatory DNA elements like enhancers. On the linear chromatin template these elements can be located hundreds of kilobases away from their target gene.
- Methods like 3C are able to map the structural organization of the chromatin fibre. Numerous 3C studies have demonstrated that the chromatin fibre of gene loci when activated folds into an intricate configuration. This organization brings the enhancers into close proximity to the promoters of the activated genes while the intervening chromatin fibre loops out.
- Transcription factors and associated protein complexes mediate the maintenance of the long-range chromatin interactions between enhancers and promoters.
- How the chromatin loops between enhancers and promoters are set up and how the proximity of an enhancer to a promoter subsequently enhances gene transcription is essentially still unknown.

Acknowledgements
I apologize to those whose work could not be discussed here due to space limitations.

FUNDING
Netherlands Organization for Scientific Research (NWO) (grant number 700.57.408).

References


121. Schubeler D, Groudine M, Bender MA. The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. *Proc Natl Acad Sci USA* 2001;98:11432–7.


