

DNA-binding factor CTCF and long-range gene interactions in V(D)J recombination and oncogene activation

*Claudia Ribeiro de Almeida,¹ *Ralph Stadhouders,² Supat Thongjuea,³ Eric Soler,^{2,4} and Rudi W. Hendriks¹

Departments of ¹Pulmonary Medicine and ²Cell Biology and Genetics, Erasmus Medical Center (MC) Rotterdam, Rotterdam, The Netherlands; ³Computational Biology Unit, Bergen Center for Computational Science, University of Bergen, Bergen, Norway; and ⁴Cancer Genomics Center, Erasmus MC Rotterdam, Rotterdam, The Netherlands

Regulation of V(D)J recombination events at immunoglobulin (Ig) and T-cell receptor loci in lymphoid cells is complex and achieved via changes in substrate accessibility. Various studies over the last year have identified the DNA-binding zinc-finger protein CCCTC-binding factor (CTCF) as a crucial regulator of long-range chromatin interactions. CTCF often controls specific interactions by preventing inappropriate communication between neighboring regulatory elements or inde-

pendent chromatin domains. Although recent gene targeting experiments demonstrated that the presence of the CTCF protein is not required for the process of V(D)J recombination per se, CTCF turned out to be essential to control order, lineage specificity and to balance the Ig V gene repertoire. Moreover, CTCF was shown to restrict activity of κ enhancer elements to the Ig κ locus. In this review, we discuss CTCF function in the regulation of V(D)J recombination on the basis

of established knowledge on CTCF-mediated chromatin loop domains in various other loci, including the imprinted H19-Igf2 locus as well as the complex β -globin, MHC class II and IFN- γ loci. Moreover, we discuss that loss of CTCF-mediated restriction of enhancer activity may well contribute to oncogenic activation, when in chromosomal translocations Ig enhancer elements and oncogenes appear in a novel genomic context. (*Blood*. 2012;119(26):6209-6218)

Introduction

Development of B and T lymphocytes requires proper spatial and temporal control of gene expression. This is achieved through the coordinate binding of lineage-specific transcription factors to *cis*-regulatory elements. In many developmentally regulated gene clusters (eg, the antigen receptor and cytokine loci), gene regulation is particularly complex and involves cell type-specific interactions over large genomic distances between gene promoters and remote enhancers, silencers, or locus control regions (LCRs).^{1,2} Regulation of higher-order chromatin structures is mediated by protein complexes composed of lineage-specific and ubiquitously expressed nuclear factors. It has recently become apparent that the DNA-binding protein CCCTC-binding factor (CTCF) is a leading candidate for the regulation of gene expression at complex loci during lymphoid differentiation.

General introduction into CTCF biology

CTCF is a ubiquitously expressed and highly conserved 11-zinc finger protein that was originally identified as a negative transcriptional regulator of the *c-Myc* oncogene.³ CTCF has since been implicated in a variety of regulatory functions, including transcriptional activation and repression, enhancer blocking, gene insulation, gene silencing, genomic imprinting, X chromosome inactivation, and long-range chromatin interactions.⁴

CTCF occupancy studies in various human and murine cell types, including embryonic stem cells and T lymphocytes, revealed approximately 14 000 to approximately 40 000 CTCF-binding sites, sharing a 15- to 20-bp core consensus sequence.⁵⁻⁸ Although

the genome-wide distribution of these sites strongly correlates with gene density, it suggests a role for CTCF distinct from that of canonical transcription factors: binding patterns do not correlate with transcriptional start sites and do not appear to predict cell-specific gene expression levels. Domains with few or no CTCF-binding sites tend to include clusters of transcriptionally coregulated genes that are often flanked by CTCF-binding sites. Conversely, CTCF-rich regions contain genes displaying extensive alternative promoter usage, including the T-cell receptor (TCR) and immunoglobulin (Ig) gene loci.⁴

Many studies have suggested that CTCF is the archetypal vertebrate protein that binds insulator elements. Insulators affect gene expression by blocking inappropriate communication between neighboring regulatory elements (enhancer-blocking activity, when positioned between enhancer and promoter) and by preventing the spread of repressive heterochromatin (barrier function).⁹ Based on studies showing that a 5' boundary element of the chicken β -globin locus serves as an insulator in erythroid cells and protects against position effect in *Drosophila*, Chung et al were the first to hypothesize that insulator function may depend on interactions between distant elements that result in looping out of intervening sequences.¹⁰ The first association of CTCF with enhancer blocking activity was proposed on the basis of CTCF-binding to the DNase hypersensitive site (5'HS4) upstream of the chicken β -globin locus and another insulator sequence at the 3' end.¹¹ Based on functional analysis in transgenic systems, CTCF has been demonstrated to mediate insulation at several model loci discussed below in detail. In addition, computational analyses showing CTCF binding to highly conserved noncoding elements

Submitted March 19, 2012; accepted April 19, 2012. Prepublished online as *Blood* First Edition paper, April 26, 2012; DOI 10.1182/blood-2012-03-402586.

*C.R.d.A. and R.S. contributed equally to this study.

© 2012 by The American Society of Hematology

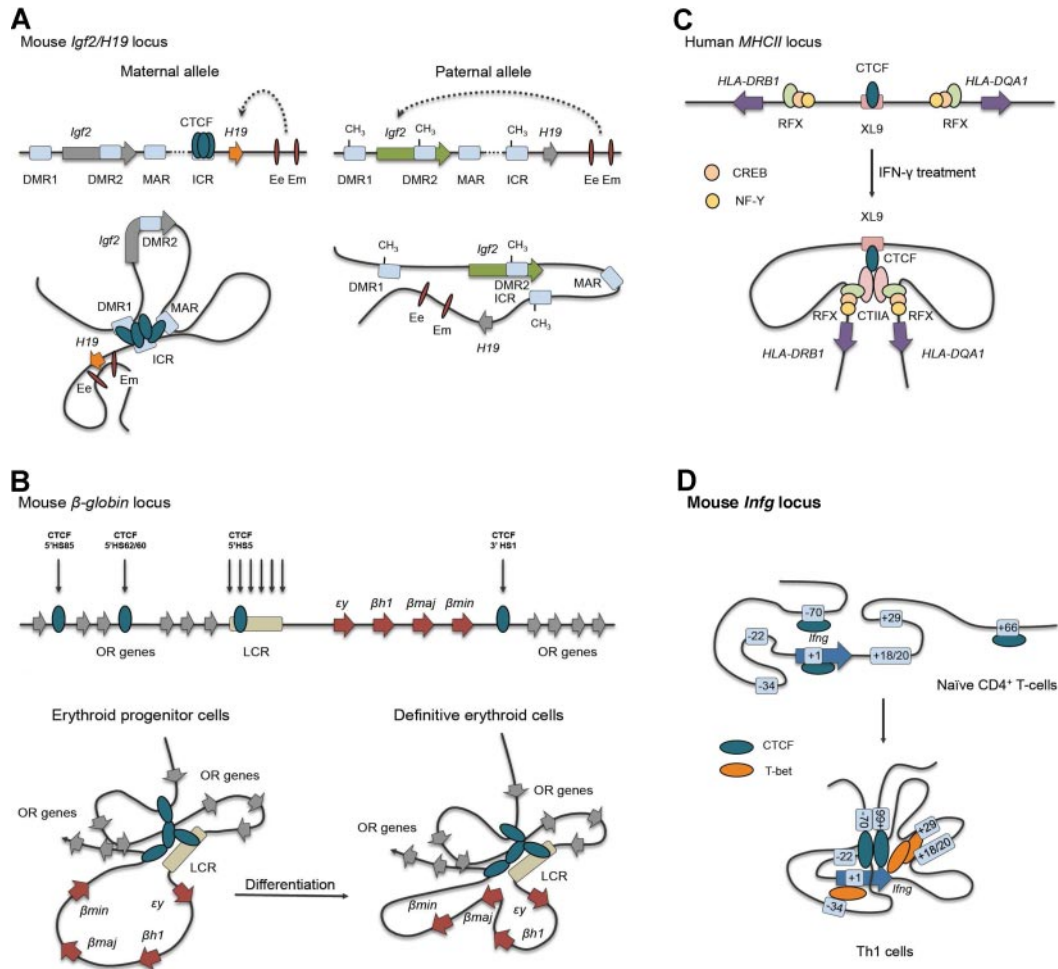


Figure 1. CTCF-mediated gene insulation via chromatin looping at several model loci. (A) Allele-specific, DNA methylation-sensitive binding of CTCF and chromatin looping account for parent-of-origin *H19* and *Igf2* gene expression patterns during mouse development. DMR indicates differentially methylated region; Ee, endodermal tissue enhancer; ICR, imprinting control region; and Em, mesodermal tissue enhancer. (B) In progenitor cells, mouse *globin* genes embedded within the olfactory receptor (OR) gene cluster are looped out from the CTCF-mediated chromatin hub, and ϵy and $\beta h1$ are expressed. On differentiation, dynamic interactions between the definitive erythroid genes β -major (βmaj) and β -minor (βmin) and the LCR are observed. HS indicates hypersensitive site. (C) IFN- γ treatment induces cotranscription of the human *HLA-DRB1* and *HLA-DQA1* genes in nonexpressing cell types by up-regulating the coactivator CIITA. Gene activation results from heterodimer formation between CIITA and the transcription factors RFX, CREB, and NF-Y bound to the inactive promoter of the *MHC class II* genes and the CTCF-bound XL9 enhancer. (D) On mouse Th1 cell differentiation, CTCF cooperates with T-bet to juxtapose distal regulatory elements, including the conserved noncoding sequence (CNS)-34 enhancer, to the *Ifng* gene. This results in a Th1-cell-specific chromatin structure at the *Ifng* locus for efficient IFN- γ transcription.

across 12 mammalian species provided evidence supporting a global role for CTCF as an insulator protein.¹²

CTCF-mediated gene insulation via chromatin loops in model loci

CTCF drives enhancer-blocking insulation through long-range chromatin interactions, in particular through 3-dimensional loop formation between distal elements.^{4,13}

The *H19/Igf2* locus

CTCF has been implicated in genomic imprinting, the epigenetically regulated process that causes genes to be expressed in a parental-origin-specific manner rather than from both chromosomes.¹⁴ In the *H19/Igf2* locus, an imprinting control region, containing 4 CTCF-binding sites, is located immediately upstream of the *H19* gene and is essential for regulation of imprinted expression of the *H19* and *Igf2* genes (Figure 1A). On the

maternally inherited allele, the imprinting control region is unmethylated, allowing for CTCF binding and the formation of a tightly coiled chromatin loop insulating the *Igf2* gene and preventing its promoter to access the enhancers downstream of *H19*. On the paternal allele, however, the imprinting control region is methylated, which abrogates CTCF binding, resulting in functional interactions between the *Igf2* promoters and the enhancer elements (Figure 1A). Moreover, loss of CTCF binding is also associated with *H19* promoter methylation leading to *H19* gene silencing on the paternal allele.

The β -globin locus

CTCF-mediated contacts are associated with transcriptional activation of genes in the extensively characterized β -globin locus in erythroid cells.¹⁵ Developmentally regulated expression of individual *globin* genes is controlled by many *cis*-acting elements, the most important of which are located in the upstream LCR (Figure 1B). CTCF binds to 3 DNase hypersensitive sites upstream in the LCR and 1 downstream (3'HS1) of the mouse β -globin locus in a

cell type–specific manner.¹⁶ During erythroid differentiation, CTCF-bound regulatory sequences throughout the locus come into close spatial proximity to form an “active chromatin hub,” as was shown by chromosome conformation capture (3C) technology detecting long-range DNA interactions.^{15–17} Formation of this “active chromatin hub” allows stage-specific interaction of active *globin* genes with the LCR and concomitant looping-out of transcriptionally silent isoforms (Figure 1B). Conditional CTCF deletion or mutation of the 3’HS1 CTCF-binding site resulted in disruption of CTCF contacts in erythroid progenitors but surprisingly had no effects on kinetics or levels of *globin* expression during erythroid differentiation.¹⁶ Because of the high density of CTCF-binding sites in the β -*globin* locus, this finding is explained by redundancy of chromatin contacts, which may not be abrogated by mutations in a single regulatory element. Genetic disruption of CTCF-binding to 3’HS1 also did not affect transcription of the neighboring olfactory receptor genes.¹⁶ Collectively, these findings indicate that correct *globin* gene expression can be achieved by different CTCF-dependent loops and thus by different chromatin conformations.

The MHC class II locus

CTCF organizes the MHC class II locus into a conformation that rearranges in the presence of the crucial non-DNA-binding transcriptional coactivator CIITA. In the human locus, CTCF binds to the intergenic enhancer XL9 that coregulates expression of the *HLA-DRB1* and *HLA-DQA1* genes^{18,19} (Figure 1C). Induction of CIITA in non-MHC class II–expressing cells results in the initiation of interactions between XL9 and the promoters upstream of *HLA-DRB1* and *HLA-DQA1*. Hereby chromatin loop formation depends on XL9-bound CTCF, CIITA, as well as the transcription factor RFX, which is constitutively bound to regulatory sequences within the *HLA-DRB1* and *HLA-DQA1* proximal promoters. Therefore, in the human MHC class II locus, CTCF is involved in the control of gene expression through loop formation via hetero-multimerization.

The *Irfng* and *Th2* cytokine loci

The key transcription factor of T helper-1 (Th1) cells, T-bet, promotes the expression of IFN- γ in part by facilitating CTCF binding in the *Irfng* locus.²⁰ CTCF acts to establish a Th1 cell-specific *Irfng* locus architecture, which helps to drive the juxtaposition of T-bet-binding enhancers, the flanking CTCF-binding elements, and CTCF sites in the first exon of the *Irfng* gene (Figure 1D). Cooperation between CTCF and T-bet is required for robust Th1 cell-specific IFN- γ expression,²⁰ which is also supported by our finding of an approximately 50% reduction in IFN- γ -producing cells when naive CTCF-deficient CD4⁺ T cells were cultured under Th1-polarized conditions.²¹ CTCF may act in a similar fashion in the Th2 cytokine locus containing the genes for IL-4, IL-5, and IL-13. Indeed, Th2 cytokines were strongly reduced in CTCF-deficient T lymphocytes cultured under Th2-polarized conditions, despite proper induction of the key Th2 transcriptional regulators Gata3 and Satb1.²¹

CTCF as a genome-wide organizer of chromatin architecture

Its capacity to form dimers and to participate in multiprotein complexes allows CTCF to organize long-range 3-dimensional looping of genomic regions. Interestingly, genome-wide identifica-

tion of binding sites for CTCF and for cohesin, the protein complex responsible for sister chromatid cohesion, showed that a large proportion of these sites are shared.^{22,23} CTCF enables cohesin recruitment to specific sites, and cohesin is required for CTCF-mediated long-range intrachromosomal interaction (eg, at the *H19/Igf2*, *Irfng*, or MHC class II loci).^{18,22,24} Because cohesin forms a large ring structure that holds sister chromatids together during mitosis, it is conceivable that cohesin may function to stabilize loops that have been initiated by CTCF.

Several lines of evidence support a role for CTCF as a crucial genome-wide organizer of chromatin architecture: (1) the high number of CTCF-binding sites in the genome, both near transcription start sites, in exons, in introns, and in intergenic regions⁴; (2) the essential function of CTCF in the orchestration of long-range chromatin interaction at model loci; (3) CTCF has many binding partners, including transcription factors, chromatin-modifying proteins, the nucleolar protein nucleophosmin, RNA polymerases, and Smad signaling proteins^{4,25}; and (4) the recent genome-wide identification of CTCF-mediated chromatin interactions implicates CTCF in the global organization of chromatin.^{26,27} Based on combining CTCF binding and mapping of histone modifications, CTCF often demarcates the boundaries of repressed chromatin associated with nuclear lamina domains.²⁶ Recent 3C-based analyses showed that indeed CTCF-mediated loops separate H3K4 active domains from H3K27me3 repressive domains, which coincide with lamina domains.²⁷ These findings demonstrate that CTCF is involved in the division of the mammalian genome into large, discrete domains that are units of chromosome organization within the nucleus.

Regulation of gene rearrangements in antigen receptor loci

Lymphocyte development is a stepwise process involving ordered rearrangement of variable (V), diversity (D), and joining (J) antigen receptor gene segments, also known as V(D)J recombination.^{1,28} Because these gene segments are spread over several megabases of the genome, it is obvious that locus compaction and looping are essential for the site-specific DNA recombination reactions that create a diverse repertoire of Ig and TCRs. The question how V(D)J recombination is regulated over great distances is, however, largely unanswered. Various studies over the last year have identified CTCF as a crucial regulator of long-range chromatin interactions in antigen receptor loci.

The lymphoid-specific recombination activation gene endonucleases Rag1 and Rag2 cooperate to induce double-strand DNA breaks at recombination signal sequences flanking the V, D, and J gene segments, which are subsequently joined by the nonhomologous end-joining DNA-repair machinery.^{1,28} Cell cycle-dependent regulation of Rag expression during lymphocyte development is critical for preventing Rag-mediated DNA breaks leading to translocation in lymphocytes undergoing active V(D)J recombination.²⁹ B-cell development is initiated in pro-B cells in the bone marrow by sequential D-to-J_H and V-to-DJ_H joining. Successful rearrangement of the Ig heavy chain (*Igh*) locus results in IgH μ protein deposition on the surface of pre-B cells, together with the preexisting surrogate light chain proteins λ 5 and VpreB, as the pre-B cell receptor (pre-BCR) complex. Pre-BCR signaling controls the selective expansion of cells that have a functionally rearranged IgH chain and redirects V(D) recombination activity from the *Igh* to the *Igk* or *Igl* light chain loci.^{30,31} Successful

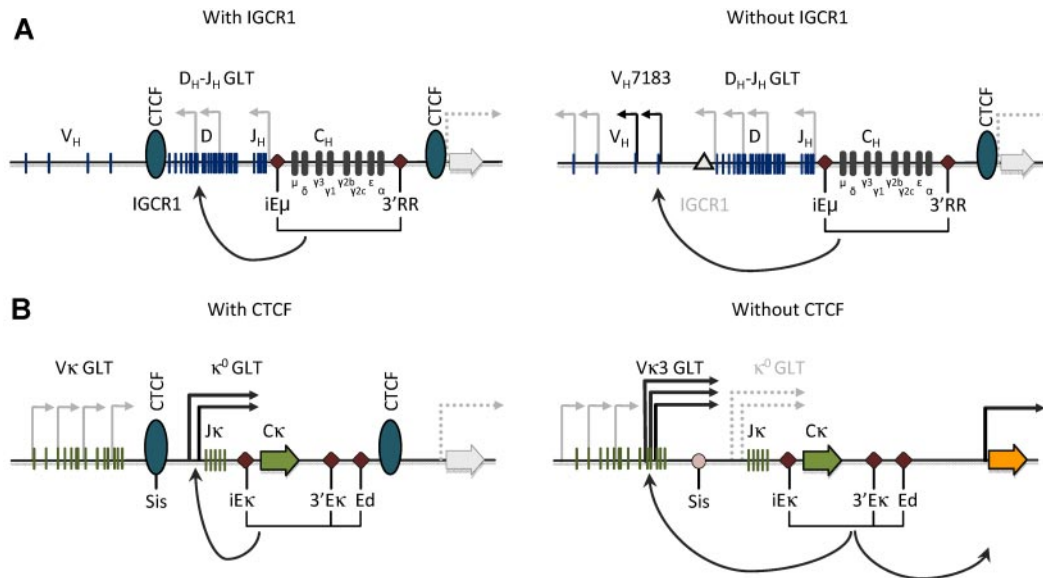


Figure 2. Model for CTCF function at the *Igh* and *Igk* loci. (A) CTCF binding at the IGCR1 in the V_H-D intergenic region in the *Igh* locus prevents inappropriate communication between the *Igh* enhancers and the proximal V_H gene segments. Targeted IGCR1 deletion in the mouse (right) leads to antisense transcription extending into proximal V_H segments of the V_H7183 family and concomitant aberrant rearrangement.^{53,54} GLT indicates germline transcript; and IGCR1, intergenic control region 1. (B) Likewise, CTCF binding at the SIS element in the V_κ-J_κ intergenic region in the *Igk* locus is important for proper germline transcription over the J_κ region, limits proximal V_κ recombination, and restricts κ enhancer interactions to the *Igk* locus.^{49,60} Conditional deletion of Ctcf in the mouse (right) leads to increased proximal V_κ-family transcription and rearrangement, and increased expression of the neighboring *Rpl3* gene (orange arrow). GLT indicates germline transcript; and SIS, silencer of intergenic sequence.

Igk/Igλ chain rearrangement results in surface expression of the BCR and additional checkpoints follow to limit autoreactivity.³² Likewise, in the thymus, V(D)J recombination events in the TCR loci accompany the development of mature T cells, expressing an αβ TCR or, in a minor subset, a γδ TCR on the cell surface.³²

Given that all *Ig* and *TCR* loci are rearranged by a common recombination machinery, an essential level of regulation of the V(D)J recombination process relies on lineage and developmental stage-specific changes in the accessibility of antigen receptor loci for recombination. The “accessibility hypothesis” was proposed based on the initial observation that rearranging gene segments are concomitantly undergoing germline transcription.^{28,33} It has been demonstrated that transcription itself can direct recombinase targeting: blockade of transcriptional elongation through the mouse TCR-α locus suppressed Vα to Jα recombination and chromatin remodeling of Jα segments.³⁴ Germline transcription increases recombination of Jα segments located close to an upstream crucial promoter, called TEA, and prevents activation of downstream promoters through transcriptional interference.³⁴ Accessibility is controlled by local *cis*-regulatory elements (promoters and enhancers) and involves subnuclear relocation, histone acetylation and/or methylation, DNA demethylation, and locus compaction.^{1,28} These processes are dependent on various lineage-specific transcription factors that bind to enhancer elements present in the *Ig* and *TCR* loci.³⁵⁻³⁷ Among these transcription factors, Pax5, EBF, and E2A play a central role in B-cell differentiation, as they also establish a hierarchical regulatory network for specification and commitment to the B-cell fate.³⁸ Next to CTCF, only a few general nuclear factors that modulate antigen receptor loci topology have been identified to date, including YY1 and Ezh2.^{35,37,39,40} However, the precise mechanism by which these factors act in concert with

lineage-specific transcription factors to regulate chromatin structure at the Ig loci remains unclear.

CTCF function in *IgH* locus rearrangement

The *Igh* locus in the mouse spans an approximately 2.7-Mbp region and is composed of more than 100 V_H gene segments and D_H, J_H, and C_H clusters located downstream. The locus contains 2 important regulatory elements: the intronic enhancer (iEμ), located 3' of the J_H genes, and the 3' regulator region (3'RR; Figure 2A).²⁸ The stepwise activation of the locus, reflected by D_H-to-J_H rearrangement, is paralleled by germline transcription. First, Iμ transcripts are generated from iEμ, μ0 transcripts from the promoter of the most 3' D_H segments, and antisense intergenic transcripts throughout the D_H and J_H domains.^{28,41,42} After D_H-to-J_H recombination, noncoding sense and antisense transcripts are generated from the V_H genes.^{28,33,42,43} Three-dimensional FISH experiments have elucidated why pro-B cells deficient for Pax5, Yy1, or Ikaros only recombine D-proximal V_H genes, suggesting that locus compaction mediated by these proteins is necessary to bring distal V_H genes in close proximity to D_H-J_H.^{35,37,39} Spatial distance measurements of genomic markers spanning the entire *Igh* locus have demonstrated that this locus is organized into compartments that contain clusters of loops or rosette-type structures separated by linkers.⁴⁴ This topology compacts the V_H region and should ensure that V_H gene segments scattered over approximately 2.5 Mbp encounter D_H-J_H segments or enhancers with appropriate frequencies.

Emerging data have revealed CTCF as an attractive candidate for controlling *Igh* locus topology. CTCF and cohesin bind to multiple sites throughout the V_H region,⁴⁵ implicating CTCF in *Igh*

chromatin loop formation and effective distal V_H - D_H - J_H recombination.^{46,47} However, CTCF occupancy in the V_H region is largely unchanged between pro-B and pre-B cells,⁴⁵ making it unlikely that CTCF alone accounts for *Igh* locus topology changes. Consistently, 3C-based experiments demonstrated that shRNA-mediated CTCF depletion only modestly affected *Igh* locus contraction.⁴⁸ In agreement with these findings, on *mb1-Cre*-mediated conditional deletion of the *Ctcf* gene in the B-cell lineage, $Ig\mu^+$ pre-B cells were still generated and both distal and proximal V_H gene segments were used.⁴⁹

Using chromatin immunoprecipitation coupled to microarray analysis (ChIP-chip) technology, Ebert et al have mapped *Igh* active histone modifications and transcription factor occupancy.⁵⁰ They hereby identified 14 copies of a potential regulatory Pax5-activated intergenic repeat (PAIR) element in the distal *Igh* region, upstream of V_H 3609 genes and interspersed among V_H 558 genes. These PAIR elements contain Pax5-dependent active chromatin and binding sites for E2A, CTCF, and the cohesin component Rad21. Detailed analysis of 3 PAIR elements showed that immediately upstream of the CTCF-binding sites noncoding antisense transcription is initiated exclusively in pro-B cells. The authors hypothesize that PAIR elements mediate looping of distal V_H genes in pro-B cells, whereby Pax5 orchestrates the formation of a bridging complex that mediates long-range interactions between distal and proximal V_H gene regions. In contrast, on the basis of anti-CTCF Chip-loop assays, Guo et al⁵¹ presented evidence for a CTCF-mediated but $iE\mu$ -independent multiloop domain of approximately 500 kb in the 5' V_H 558/ V_H 3609 region containing the PAIR elements, whereby locus contraction may result from interaction between $E\mu$ -bound³⁹ Yy1 and Yy1 sites in the distal V_H region. These interactions may be increased by heterotypic interactions with CTCF sites or by inter-PAIR interactions in the distal V_H region.⁵¹ Future experiments should focus on functional characterization of the PAIR elements, which would require deletion of these proposed regulatory elements to test the effects on V(D)J recombination and V_H repertoire in vivo.

In addition, CTCF-binding sites are present in the intergenic control region I (IGCR1), which is located 3 to 5 kb upstream of the most 5' D_H element (Figure 2A).^{45,52,53} These CTCF-binding sites have enhancer-blocking activity and mark a sharp decrease in $E\mu$ -dependent antisense D_H - J_H transcription.⁵² Targeted IGCR1 deletion in the mouse leads to D_H - J_H antisense transcripts extending into proximal V_H segments and concomitant aberrant rearrangement of the proximal V_H gene segments in developing thymocytes.^{53,54} Importantly, IGCR1 balances proximal versus distal V_H usage by inhibiting germline transcription and rearrangement of D-proximal V_H 7183 gene segments (Figure 2A). Moreover, IGCR1 maintains ordered V(D)J recombination by suppressing V_H joining to D segments that are not joined to J_H . Both FISH and 3C-based experiments showed that CTCF-binding sites in IGCR1 are in close spatial proximity with a region immediately downstream of 3'RR with a high density of CTCF-binding sites (Figure 2A).^{48,51,55} These CTCF-mediated interactions would create a domain that contains all of the D_H , J_H , and C region elements, as well as $E\mu$ and 3'RR. It therefore seems probable that CTCF binding to IGCR1 prevents chromatin accessibility extending into the independently activated V_H chromatin domains. This notion is supported by the findings that (1) insertion of a V_H gene segment just 5' of the D_H cluster resulted in loss of ordered and lineage-specific *Igh* locus rearrangement,⁵⁶ and (2) mice with an approximately 8-kb deletion encompassing the CTCF-binding sites downstream 3'RR showed a 2-fold increase in the usage of V_H 7183.⁵⁷

In summary, by binding to IGCR1 and downstream of 3'RR, CTCF acts as an insulator that ensures ordered and lineage-specific V(D)J recombination at the *Igh* locus.

CTCF function in *Igk* locus gene rearrangement

The *Igk* locus in the mouse spans approximately 3.2 Mbp and contains approximately 100 V_k gene segments, a J_k cluster and a single C_k region.²⁸ *Igk* locus rearrangement and expression are positively regulated by an intronic enhancer (iE κ) located between J_k and C_k , and 2 enhancers downstream of C_k (3'E κ and Ed; Figure 2B). As D elements are lacking, κ light chains can be expressed on a single productive V_k -to- J_k recombination event and autoreactive BCR specificities can be replaced by "receptor editing," the process of ongoing Ig light chain gene recombination.^{1,28} Germline κ^0 transcripts from promoters located upstream of J_k (κ^0 1.1 and κ^0 0.8) or V_k gene segments indicate the accessibility over J_k and V_k regions, respectively.⁵⁸ A recombination silencer element in the V_k - J_k region, called "silencer in intervening sequence" (SIS), negatively regulates *Igk* rearrangement by targeting of the nonrearranging allele to centromeric heterochromatin (Figure 2B).⁵⁹ Deletion of the SIS element resulted in decreased monoallelic *Igk* heterochromatin localization, reduced κ^0 germline transcription, and enhanced proximal V_k usage.^{59,60}

Next to strong CTCF binding at the 5' and 3' boundaries and the SIS element,⁴⁹ we identified approximately 60 CTCF-binding sites in the V_k region (in contrast to the previously reported low density of CTCF occupancy in the *Igk* locus⁴⁵). CTCF-binding sites were not evenly distributed over the *Igk* locus. Detailed analysis of long-range interactions in pre-B cells revealed a correlation between V_k usage, local density of CTCF-binding sites, and the frequencies of contacts with SIS, iE κ , or 3'E κ .⁴⁹ Using 3C-seq, we observed that long-range interactions between the SIS-, iE κ -, or 3'E κ - and V_k region were not uniformly affected by conditional deletion of the *CTCF* gene in the B-cell lineage. Loss of CTCF resulted in significantly increased interactions with the most proximal V_k region containing the V_k 3 family. In line with these findings, CTCF-deficient pre-B cells showed significantly increased V_k 3 germline transcription and recombination.⁴⁹ On the other hand, loss of CTCF resulted in reduced contacts with the 2 most distal V_k genes, V_k 2-137 and V_k 1-135, which are often rearranged in control but not in CTCF-deficient (pre-)B cells.

Based on these studies,^{49,59,60} we proposed that CTCF binding in the SIS element provides enhancer-blocking activity that limits interactions between the enhancers and the proximal V_k promoters. CTCF, however, plays a dual role, as, next to the enhancer-blocking activity, it tethers remote V_k gene promoters to the enhancers.

Comparison of CTCF function in the *Igh* and *Igk* locus

The enhancer blocking activity of CTCF at the SIS element in the V_k - J_k intergenic region in the *Igk* locus^{49,60} remarkably parallels the role of CTCF at IGCR1 in the V_H -D intergenic region in the *Igh* locus.⁵³ Likewise, in thymocytes, CTCF colocalizes with cohesin at the TCR- α regulatory elements TEA just upstream of J_α . Loss of cohesin in double-positive

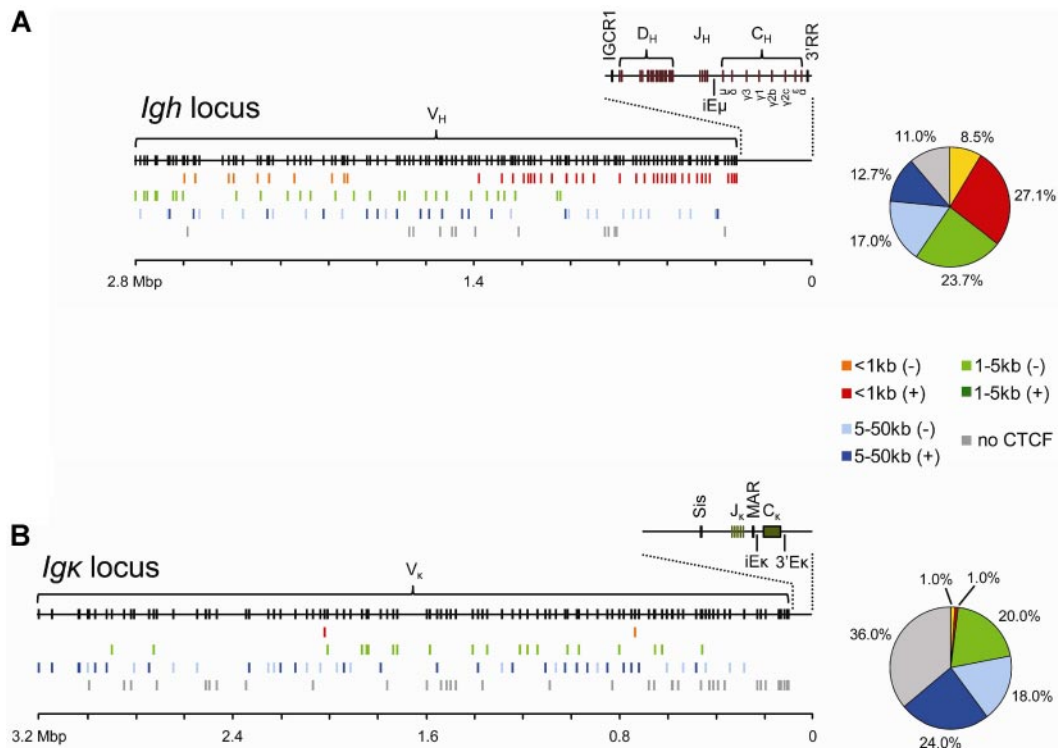


Figure 3. V_H and V_κ gene positions relative to CTCF occupancy. (A) Overview of V_H gene positions relative to the nearest CTCF binding site for the mouse *Igh* locus. The germline *Igh* locus is shown with the locations of V_H , D_H , J_H , C_H , and enhancer elements (top). V_H genes were grouped according to the location of the nearest CTCF site and whether this site was located upstream (-) or downstream (+) relative to V_H transcriptional orientation. Color-coded groups of V_H genes were plotted at their respective locations below the germline V_H genes. On the right, a pie-chart represents the distribution of the different V_H gene groups expressed in percentages of the total number of V_H genes. (B) Identical to panel A, but for the V_κ genes within the mouse *Igk* locus.

thymocytes led to reduced long-range TEA-E α interactions and significantly decreased germline transcription and rearrangement over the 3'J α gene segments.⁶¹ Cohesin and CTCF also colocalize at 5' boundary sites just downstream of the E α enhancer, demarcating the *TCR- α* locus from the neighboring gene *Dad1*,⁶¹ a region with CTCF-dependent insulator activity.⁶²

On virtually complete deletion of CTCF protein in the B-cell lineage in our conditionally targeted mice, we found that productive *Igh* and *Igk* rearrangements were still generated, demonstrating that CTCF is not required for V(D)J recombination per se. This is also consistent with the presence of TCR- α - or TCR- β -chain rearrangement in CTCF-deficient thymocytes in vivo.⁶³ The *Igh* and *Igk* loci both contain approximately 60 CTCF-binding sites that influence the genomic architecture of these loci. The identification in the *Igh* locus of PAIR elements just upstream of distal V_H 3609 genes,⁵⁰ together with the observation that many proximal V_H gene segments are located within 100 bp of CTCF-binding sites,⁶⁴ suggests that proximity of a CTCF-binding site may affect the probabilities of individual V_H genes to encounter a D_H - J_H element for recombination. In this aspect, the *Igh* and *Igk* loci are remarkably different: whereas approximately 27% of V_H genes have a CTCF-binding site within 1-kb distance, this is the case for only 2 V_κ genes (Figure 3). Moreover, only approximately 11% of V_H genes but approximately 36% V_κ genes do not have a CTCF-binding site in the intervening sequences up to their 5' or 3' neighboring V genes. Thus, although CTCF binds at many sites throughout the *Igk* locus, most CTCF occupancy is located between V_κ segments and not adjacent to the V_κ segments.^{1,49} It has been

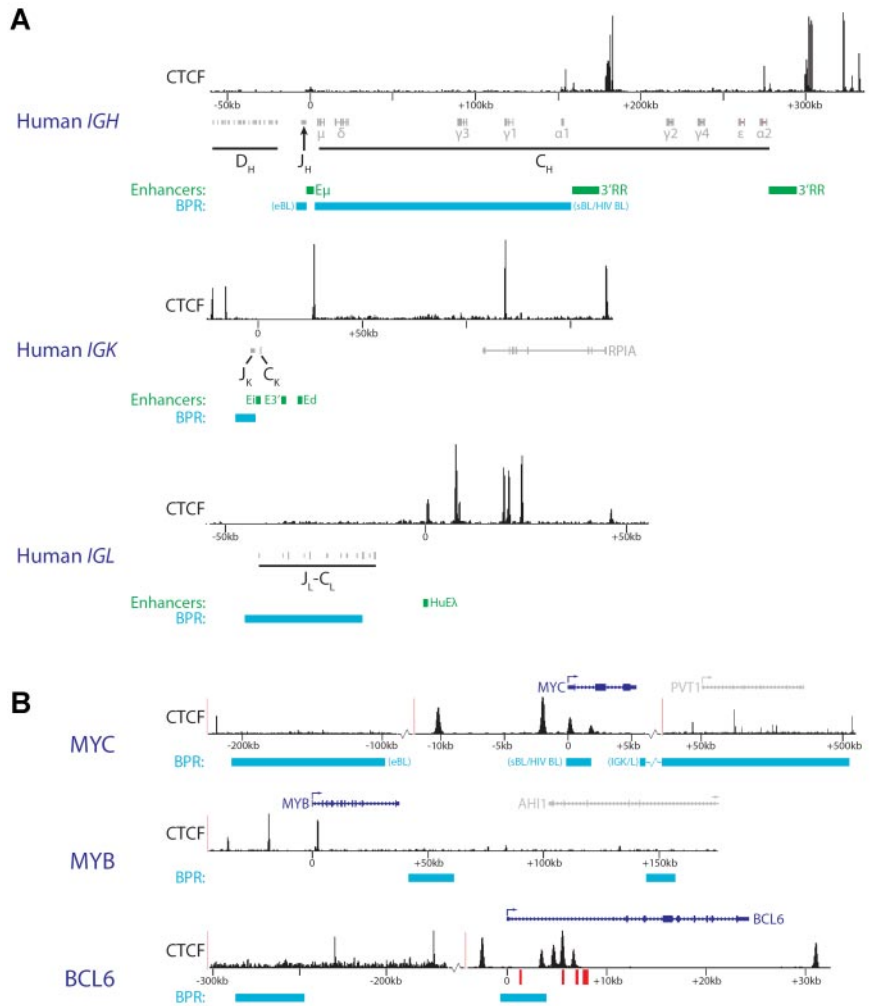
suggested that V regions are organized as rosettes by CTCF, whereby CTCF-binding adjacent to a V gene increases its recombination probability.⁶⁴ Because the *Igh* and *Igk* loci show large differences in the proximity of CTCF-binding sites to V genes, in such a model the 2 loci would be very differently organized in 3-dimensional space to provide appropriate access of individual V regions to the proposed recombination center.⁶⁵ Interestingly, it was recently proposed that E2A proteins may modulate *Igk* locus topology by acting as anchors in a mechanism similar to that put forward for CTCF in the *Igh* locus.¹ This hypothesis still needs to be tested but is supported by the remarkable, nonrandom distribution of E2A binding sites across the *Igk* locus, in particular within 200 bp of the 5' or the 3' end of V_κ regions.^{1,46,47}

CTCF and long-range gene interactions in Ig enhancer-mediated oncogene activation

Many leukemias and lymphomas contain chromosomal translocation involving Ig or TCR loci whereby the strong enhancers in these loci are placed in a new genomic context. Because CTCF and cohesin have been implicated in interchromosomal interactions, they probably contribute to physical proximity of particular translocation targets in lymphocytes. For example, in many Burkitt B-cell lymphomas, translocations place the *MYC* oncogene up to approximately 200 kb upstream of the *IgH* 3'RR (Figure 4A). In normal B cells, *IgH* and *MYC* are preferentially positioned in close spatial proximity relative to each other and

Figure 4. CTCF distribution and chromosomal breakpoint regions (BPR) around Ig enhancers and oncogenes involved in translocation with antigen receptor loci.

(A) Genomic regions surrounding the human IgH enhancers. CTCF ChIP-Seq data (from the human GM12878 B lymphoblastoid cell line⁸¹) is shown above the localization of D_H, J_H, and C_H fragments. Green represents enhancers; and cyan, most common chromosomal BPRs. For the *IGH* locus, 2 common BPRs are annotated. eBL indicates endemic Burkitt lymphoma; sBL, sporadic Burkitt lymphoma; and HIV BL, HIV-induced Burkitt lymphoma. For scaling purposes, the position of E_μ (*IGH*), E_i (*IGK*), or HuE_λ (*IGL*) enhancers was set to 0. (B) CTCF ChIP-Seq data (from the human GM12878 B lymphoblastoid cell line⁸¹) are shown for 3 genomic regions around oncogenes involved in translocations with antigen receptor loci. Locations of genes are shown on top, with the relevant oncogenes in blue and other genes in gray. The oncogene transcriptional start site was set to 0 for scaling purposes. Cyan indicates the most common chromosomal BPRs. BPRs linked to distinct B-cell malignancies are indicated as in panel A. For *MYC*, breakpoints found specifically in translocations with Ig light chain loci are annotated (*IGK/L*). Red bars represent the *BCL6* CpG islands located in intron 1.



are present in the same transcription factories.⁶⁶⁻⁶⁸ Recently developed new genome-wide methods that measure primary chromosomal rearrangements in the absence of growth selection have shown that translocations between *IGH* and *MYC* and all other genes are indeed directly related to their contact frequencies.⁶⁹⁻⁷² However, 3C-based measurements showed that approximately 30% of all genes interact with *IGH* at equal or higher frequency than *MYC*.⁷¹ Therefore, growth selection and AID-mediated DNA damage rather than a high contact frequency between *IGH* and the translocations partners account for their high rate of translocation.

An oncogenic role of the *IGH* locus 3'RR has been defined based on long-range activation of translocated *MYC* genes.⁷³ In this context, an important function of CTCF is to restrict enhancer interaction to Ig loci⁴⁹ (Figure 4). Obviously, CTCF-mediated restriction might be lost when strong Ig enhancers appear in an aberrant genomic context in chromosomal translocations, which may then lead to activation of oncogenes. In addition, CTCF sites clustered around Ig enhancers may affect local chromatin conformation to promote novel promoter-enhancer interactions. The genes encoding c-Myb and c-Myc, 2 essential hematopoietic regulators and oncogenic transcription factors, harbor multiple CTCF-binding sites in their vicinity (Figure 4B). Indeed, CTCF was originally discovered as a transcriptional repressor of *c-Myc*.⁷⁴ Both genes are regulated at

the level of transcriptional elongation by attenuator elements located in their first introns,^{75,76} close to CTCF-binding sites^{77,78} (Figure 4B). The exact role played by CTCF at these sites is still unclear, but the recent demonstration that CTCF regulates RNA polymerase II processing⁷⁹ suggests that CTCF could directly affect elongation. Loss of intronic CTCF binding resulting from translocations could disrupt normal locus structure and transcriptional control, allowing the juxtaposed *IGH* enhancers to drive *MYC* expression. Because *MYB* seems to be regulated by a similar mechanism, this may also occur within the *MYB* locus, when translocated to the *TCR-β* locus in acute T-cell leukemia.⁸⁰

Interestingly, based on published CTCF-binding sites,⁸¹ translocation breakpoints at the *CRLF2* and *BCL3* loci found in several types of B-cell malignancies are both confined to an approximately 25-kb region between the transcriptional start site and an upstream CTCF element.^{82,83} Translocations therefore result in loss of the upstream CTCF-binding site, which may allow strong Ig enhancers to interact and activate the *CRLF2* and *BCL3* promoters. In contrast, many other loci frequently involved in leukemic translocations (eg, *BCL2* and *MYC*),^{84,85} show more complex breakpoint patterns and harbor numerous CTCF-binding sites (Figure 4B).

Expression of *BCL6* in diffuse large B cell lymphomas is maintained through hypermethylation of intragenic CpG islands in

its first intron.⁸⁶ Hypermethylation, which has been observed in a wide range of cancer cells, prevents CTCF-mediated silencing of *BCL6* because CTCF binds these CpG islands in a methylation-sensitive fashion. Furthermore, long-range regulatory regions approximately 150 to 260 kb upstream have been implicated in *BCL6* regulation,⁸⁷ possibly representing another layer of control in which CTCF might play an important role (Figure 4B).

In conclusion, based on studies in the *Igh* and *Igκ*, it has become clear that CTCF-binding provides enhancer-blocking activity that restricts Ig enhancer activity to the Ig loci. Nevertheless, CTCF also serves to bring remote V gene promoters into close proximity of Ig enhancers. It will be challenging to identify the exact mechanism by which CTCF regulates V gene choice because individual developing B cells are expected to manifest distinct and highly dynamic interactions between V regions and Ig enhancers. Moreover, the high density of CTCF-binding sites and their anticipated partial redundancy may complicate the interpretation of *in vivo* targeted deletion experiments.

Chromosomal translocations present in lymphoid malignancies will (1) remove CTCF sites close to the oncogenes, (2) bring in enhancer-associated CTCF sites, or (3) dramatically alter the spacing between CTCF sites and their cognate genes. These events are likely to influence local communication between *cis*-regulatory elements by imposing new structural constraints and thereby increasing oncogene expression. Additional epigenetic changes, including DNA methylation and histone modifications, might also interfere with CTCF function. For example, the *BCL6* oncogene is involved in chromosomal translocations through breakpoints frequently found near the intronic CpG islands where CTCF binds in a methylation-sensitive manner.^{86,88}

In a normal context, oncogene expression is tightly regulated in response to extracellular signaling events. It is an intriguing question why these signals are no longer able to efficiently regulate expression of the translocation products. Trying to address this issue might reveal additional roles for CTCF in the integration of

signaling to chromatin. Future research should clarify the complex roles played by CTCF in the wiring of *cis*-DNA regulatory elements and in organizing networks of genomic interactions. The field of spatial chromatin organization research is currently exploding with the development of new exciting technologies, allowing high-throughput analysis of chromatin organization of the whole genome.^{27,89,90} As the resolution of these technologies increases, they might rapidly become valuable tools to reveal the complex 3-dimensional changes in cancer genomes and unravel their relationships with critical regulatory factors, including CTCF.

Acknowledgments

This work was supported in part by Fundação para a Ciência e a Tecnologia (C.R.d.A.), Royal Netherlands Academy of Arts and Sciences (R.S.), the Center of Biomedical Genetics, the Cancer Genomics Center, and the EuTRACC Consortium (E.S. and R.S.).

Authorship

Contribution: C.R.d.A. wrote the manuscript and designed figures; R.S. analyzed and interpreted data, wrote the manuscript, and designed figures; S.T. analyzed data; E.S. wrote the manuscript; and R.W.H. designed figures and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

The current affiliation for C.R.d.A. is Genome Function Group, Medical Research Council Clinical Sciences Centre, Imperial College School of Medicine, London, United Kingdom.

Correspondence: Rudi W. Hendriks, Department of Pulmonary Medicine, Room Ee2251a, Erasmus MC Rotterdam, PO Box 2040, NL 3000 CA Rotterdam, The Netherlands; e-mail: r.hendriks@erasmusmc.nl.

References

- Bossen C, Mansson R, Murre C. Chromatin topology and the regulation of antigen receptor assembly. *Annu Rev Immunol*. 2012;30:337-356.
- Kanno Y, Vahedi G, Hirahara K, Singleton K, O'Shea JJ. Transcriptional and epigenetic control of T helper cell specification: molecular mechanisms underlying commitment and plasticity. *Annu Rev Immunol*. 2012;30:707-731.
- Lobanenkov VV, Nicolas RH, Adler VV, et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*. 1990;5(12):1743-1753.
- Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell*. 2009;137(7):1194-1211.
- Kim TH, Abdullaev ZK, Smith AD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*. 2007;128(6):1231-1245.
- Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823-837.
- Chen X, Xu H, Yuan P, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*. 2008;133(6):1106-1117.
- Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res*. 2009;19(1):24-32.
- Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Curr Opin Genet Dev*. 2007;17(5):400-407.
- Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*. 1993;74(3):505-514.
- Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*. 1999;98(3):387-396.
- Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES. Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci U S A*. 2007;104(17):7145-7150.
- Hou C, Zhao H, Tanimoto K, Dean A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A*. 2008;105(51):20398-20403.
- Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet*. 2011;12(8):565-575.
- Palstra RJ, de Laat W, Grosveld F. Beta-globin regulation and long-range interactions. *Adv Genet*. 2008;61:107-142.
- Splinter E, Heath H, Kooren J, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev*. 2006;20(17):2349-2354.
- Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell*. 2002;10(6):1453-1465.
- Majumder P, Boss JM. CTCF controls expression and chromatin architecture of the human major histocompatibility complex class II locus. *Mol Cell Biol*. 2010;30(17):4211-4223.
- Majumder P, Gomez JA, Chadwick BP, Boss JM. The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J Exp Med*. 2008;205(4):785-798.
- Sekimata M, Perez-Melgosa M, Miller SA, et al. CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-gamma locus. *Immunity*. 2009;31(4):551-564.
- Ribeiro de Almeida C, Heath H, Krpic S, et al. Critical role for the transcription regulator CCCTC-binding factor in the control of Th2 cytokine expression. *J Immunol*. 2009;182(2):999-1010.
- Wendt KS, Yoshida K, Itoh T, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008;451(7180):796-801.
- Parelho V, Hadjir S, Spivakov M, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*. 2008;132(3):422-433.
- Hadjir S, Williams LM, Ryan NK, et al. Cohesins

- form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*. 2009;460(7253):410-413.
25. Weth O, Renkawitz R. CTCF function is modulated by neighboring DNA binding factors. *Biochem Cell Biol*. 2011;89(5):459-468.
 26. Guelen L, Pagie L, Brassat E, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008;453(7197):948-951.
 27. Handoko L, Xu H, Li G, et al. CTCF-mediated functional chromatin interactions in pluripotent cells. *Nat Genet*. 2011;43(7):630-638.
 28. Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol*. 2006;24:541-570.
 29. Li Z, Dordai DI, Lee J, Desiderio S. A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. *Immunity*. 1996;5(6):575-589.
 30. Hendriks RW, Middendorp S. The pre-BCR checkpoint as a cell-autonomous proliferation switch. *Trends Immunol*. 2004;25(5):249-256.
 31. Melchers F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nat Rev Immunol*. 2005;5(7):578-584.
 32. von Boehmer H, Melchers F. Checkpoints in lymphocyte development and autoimmune disease. *Nat Immunol*. 2010;11(1):14-20.
 33. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell*. 1985;40(2):271-281.
 34. Abarrategui I, Krangel MS. Regulation of T cell receptor-alpha gene recombination by transcription. *Nat Immunol*. 2006;7(10):1109-1115.
 35. Fuxa M, Skok J, Souabni A, Salvaggio G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev*. 2004;18(4):411-422.
 36. Romanow WJ, Langerak AW, Goebel P, et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. *Mol Cell*. 2000;5(2):343-353.
 37. Reynaud D, Demarco IA, Reddy KL, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol*. 2008;9(8):927-936.
 38. Medina KL, Pongubala JM, Reddy KL, et al. Assembling a gene regulatory network for specification of the B cell fate. *Dev Cell*. 2004;7(4):607-617.
 39. Liu H, Schmidt-Suppran M, Shi Y, et al. Yin Yang 1 is a critical regulator of B-cell development. *Genes Dev*. 2007;21(10):1179-1189.
 40. Su IH, Basavaraj A, Krutchinsky AN, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol*. 2003;4(2):124-131.
 41. Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. *Nature*. 1985;318(6045):475-478.
 42. Bolland DJ, Wood AL, Johnston CM, et al. Antisense intergenic transcription in V(D)J recombination. *Nat Immunol*. 2004;5(6):630-637.
 43. Haines BB, Brodeur PH. Accessibility changes across the mouse Igh-V locus during B cell development. *Eur J Immunol*. 1998;28(12):4228-4235.
 44. Jhunjunwala S, van Zelm MC, Peak MM, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell*. 2008;133(2):265-279.
 45. Degner SC, Wong TP, Jankevicius G, Feeney AJ. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol*. 2009;182(1):44-48.
 46. Jhunjunwala S, van Zelm MC, Peak MM, Murre C. Chromatin architecture and the generation of antigen receptor diversity. *Cell*. 2009;138(3):435-448.
 47. Lin YC, Jhunjunwala S, Benner C, et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat Immunol*. 2010;11(7):635-643.
 48. Degner SC, Verma-Gaur J, Wong TP, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A*. 2011;108(23):9566-9571.
 49. Ribeiro de Almeida C, Stadhouders R, de Bruijn MJ, et al. The DNA-binding protein CTCF limits proximal V kappa recombination and restricts kappa enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity*. 2011;35(4):501-513.
 50. Ebert A, McManus S, Tagoh H, et al. The distal V(H) gene cluster of the Igh locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. *Immunity*. 2011;34(2):175-187.
 51. Guo C, Gerasimova T, Hao H, et al. Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell*. 2011;147(2):332-343.
 52. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J Biol Chem*. 2005;280(13):9327-9338.
 53. Guo C, Yoon HS, Franklin A, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature*. 2011;477(7365):424-430.
 54. Giallourakis CC, Franklin A, Guo C, et al. Elements between the Igh variable (V) and diversity (D) clusters influence antisense transcription and lineage-specific V(D)J recombination. *Proc Natl Acad Sci U S A*. 2010;107(51):22207-22212.
 55. Garrett FE, Emelyanov AV, Sepulveda MA, et al. Chromatin architecture near a potential 3' end of the Igh locus involves modular regulation of histone modifications during B-cell development and in vivo occupancy at CTCF sites. *Mol Cell Biol*. 2005;25(4):1511-1525.
 56. Bates JG, Cado D, Nolla H, Schlissel MS. Chromosomal position of a VH gene segment determines its activation and inactivation as a substrate for V(D)J recombination. *J Exp Med*. 2007;204(13):3247-3256.
 57. Volpi SA, Verma-Gaur J, Hassan R, et al. Germ-line deletion of Igh 3' regulatory region elements hs 5, 6, 7 (hs5-7) affects B cell-specific regulation, rearrangement, and insulation of the Igh locus. *J Immunol*. 2012;188(6):2556-2566.
 58. Schlissel MS, Baltimore D. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. *Cell*. 1989;58(5):1001-1007.
 59. Liu Z, Widlak P, Zou Y, et al. A recombination silencer that specifies heterochromatin positioning and Ikaros association in the immunoglobulin kappa locus. *Immunity*. 2006;24(4):405-415.
 60. Xiang Y, Zhou X, Hewitt SL, Skok JA, Garrard WT. A multifunctional element in the mouse Igh kappa locus that specifies repertoire and Ig loci subnuclear location. *J Immunol*. 2011;186(9):5356-5366.
 61. Seitan VC, Hao B, Tachibana-Konwalski K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature*. 2011;476(7361):467-471.
 62. Magdinier F, Yusufzai TM, Felsenfeld G. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *J Biol Chem*. 2004;279(24):25381-25389.
 63. Heath H, Ribeiro de Almeida C, Sleutels F, et al. CTCF regulates cell cycle progression of alpha-beta T cells in the thymus. *EMBO J*. 2008;27(21):2839-2850.
 64. Lucas JS, Bossen C, Murre C. Transcription and recombination factories: common features? *Curr Opin Cell Biol*. 2011;23(3):318-324.
 65. Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell*. 2010;141(3):419-431.
 66. Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet*. 2003;34(3):287-291.
 67. Parada LA, Sotiriou S, Misteli T. Spatial genome organization. *Exp Cell Res*. 2004;296(1):64-70.
 68. Osborne CS, Chakalova L, Mitchell JA, et al. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol*. 2007;5(8):e192.
 69. Chiarle R, Zhang Y, Frock RL, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell*. 2011;147(1):107-119.
 70. Klein IA, Resch W, Jankovic M, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell*. 2011;147(1):95-106.
 71. Hakim O, Resch W, Yamane A, et al. DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature*. 2012;484(7392):69-74.
 72. Zhang Y, McCord RP, Ho YJ, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*. 2012;148(5):908-921.
 73. Gostissa M, Yan CT, Bianco JM, Cogne M, Pinaud E, Alt FW. Long-range oncogenic activation of Igh-c-myc translocations by the Igh 3' regulatory region. *Nature*. 2009;462(7274):803-807.
 74. Klenova EM, Nicolas RH, Paterson HF, et al. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol*. 1993;13(12):7612-7624.
 75. Bentley DL, Groudine M. A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature*. 1986;321(6071):702-706.
 76. Bender TP, Thompson CB, Kuehl WM. Differential expression of c-myc mRNA in murine B lymphomas by a block to transcription elongation. *Science*. 1987;237(4821):1473-1476.
 77. Filippova GN, Fagerlie S, Klenova EM, et al. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol*. 1996;16(6):2802-2813.
 78. Stadhouders R, Thongjuea S, Andrieu-Soler C, et al. Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. *EMBO J*. 2011;31(4):986-999.
 79. Shukla S, Kavak E, Gregory M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*. 2011;479(7371):74-79.
 80. Clappier E, Cucchini W, Kalota A, et al. The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. *Blood*. 2007;110(4):1251-1261.
 81. Rosenbloom KR, Dreszer TR, Pheasant M, et al. ENCODE whole-genome data in the UCSC Genome Browser. *Nucleic Acids Res*. 2010;38(Database issue):D620-D625.

82. Dyer MJ, Akasaka T, Capasso M, et al. Immunoglobulin heavy chain locus chromosomal translocations in B-cell precursor acute lymphoblastic leukemia: rare clinical curios or potent genetic drivers? *Blood*. 2010;115(8):1490-1499.
83. McKeithan TW, Takimoto GS, Ohno H, et al. BCL3 rearrangements and t(14;19) in chronic lymphocytic leukemia and other B-cell malignancies: a molecular and cytogenetic study. *Genes Chromosomes Cancer*. 1997;20(1):64-72.
84. Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. *Blood*. 2004;104(10):3009-3020.
85. Ngan BY, Nourse J, Cleary ML. Detection of chromosomal translocation t(14;18) within the minor cluster region of bcl-2 by polymerase chain reaction and direct genomic sequencing of the enzymatically amplified DNA in follicular lymphomas. *Blood*. 1989;73(7):1759-1762.
86. Lai AY, Fatemi M, Dhasarathy A, et al. DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *J Exp Med*. 2010;207(9):1939-1950.
87. Ramachandrareddy H, Bouska A, Shen Y, et al. BCL6 promoter interacts with far upstream sequences with greatly enhanced activating histone modifications in germinal center B cells. *Proc Natl Acad Sci U S A*. 2010;107(26):11930-11935.
88. Chen YW, Hu XT, Liang AC, et al. High BCL6 expression predicts better prognosis, independent of BCL6 translocation status, translocation partner, or BCL6-deregulating mutations, in gastric lymphoma. *Blood*. 2006;108(7):2373-2383.
89. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289-293.
90. Sexton T, Yaffe E, Kenigsberg E, et al. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*. 2012;148(3):458-472.