Noncoding transcription within the *Igh* distal V_H region at PAIR elements affects the 3D structure of the *Igh* locus in pro-B cells

Jiyoti Verma-Gaur[a], Ali Torkamani[b], Lana Schaffer[c], Steven R. Head[c], Nicholas J. Schork[b], and Ann J. Feeney[a,1]

[a]Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA 92037; [b]Department of Molecular and Experimental Medicine, The Scripps Research Institute, and The Scripps Translational Science Institute, La Jolla, CA 92037; and [c]Next Generation Sequencing Core Facility, The Scripps Research Institute, La Jolla, CA 92037

Noncoding sense and antisense germ-line transcription within the Ig heavy chain locus precedes V(D)J recombination and has been proposed to be associated with *Igh* locus accessibility, although its precise role remains elusive. However, no global analysis of germ-line transcription throughout the *Igh* locus has been done. Therefore, we performed directional RNA-seq, demonstrating the locations and extent of both sense and antisense transcription throughout the *Igh* locus. Surprisingly, the majority of antisense transcripts are localized around two Pax5-activated intergenic repeat (PAIR) elements in the distal *IghV* region. Importantly, long-distance loops measured by chromosome conformation capture (3C) are observed between these two active PAIR promoters and Eμ, the start site of Iμ germ-line transcription, in a lineage- and stage-specific manner, even though this antisense transcription is Eμ-independent. $YY1^{-/-}$ pro-B cells are greatly impaired in distal V_H gene rearrangement and *Igh* locus compaction, and we demonstrate that YY1 deficiency greatly reduces antisense transcription and PAIR-Eμ interactions. ChIP-seq shows high level YY1 binding only at Eμ, but low levels near some antisense promoters. PAIR–Eμ interactions are not disrupted by DRB, which blocks transcription elongation without disrupting transcription factories once they are established, but the looping is reduced after heat-shock treatment, which disrupts transcription factories. We propose that transcription-mediated interactions, most likely at transcription factories, initially compact the *Igh* locus, bringing distal V_H genes close to the DJ_H rearrangement which is adjacent to Eμ. Therefore, we hypothesize that one key role of noncoding germ-line transcription is to facilitate locus compaction, allowing distal V_H genes to undergo efficient rearrangement.

noncoding RNA | V(D)J rearrangement | CTCF | long-range interactions

Antigen receptors in lymphocytes are assembled in the highly regulated lineage-specific process of V(D)J recombination, which creates a diverse repertoire of Ig and T-cell receptors. In each precursor lymphocyte, one each of the many V, D, and J gene segments at the appropriate receptor loci are juxtaposed to create a V(D)J exon encoding the variable region of the antigen receptor. In B-lineage progenitors, rearrangement occurs first at the Ig heavy chain (*Igh*) locus, where D_H to J_H rearrangement occurs first on both alleles, followed by V_H to DJ_H rearrangement (1). After successful rearrangement of the *Igh* locus, rearrangement at the *Igk* light chain locus begins. These successive stages of rearrangement have been proposed to be regulated by differential accessibility of different portions of the loci at the appropriate time for rearrangement (2). Early indications of this stage-specific and lineage-specific accessibility came from the observation that unrearranged gene segments underwent noncoding transcription at the stage immediately preceding their rearrangement (3, 4).

The murine *Igh* locus spans ~2.8 Mb, of which ~2.4 Mb contains the 195 V_H gene segments (5). The V_H genes are divided into 16 V_H gene families based on sequence homology. The largest V_H family is J558, which occupies the 5′ distal half of the V_H region. In order for V_H genes spread throughout this large region to find the single rearranged DJ_H gene for rearrangement, the *Igh* locus undergoes dramatic changes in its 3D structure (6–9). Three-dimensional FISH studies showed that the *Igh* locus compacts at the pro–B-cell stage, and decontracts at the subsequent pre–B-cell stage (10). By measuring spatial distances between 11 small probes spread throughout the *Igh* locus, Jhunjhunwala et al. demonstrated that distal V_H genes moved dramatically closer to the D_H genes specifically at the pro–B-cell stage (9). Computational as well as geometrical approaches in their study suggested that the *Igh* locus is organized into rosette-like clusters of loops in pre-pro-B cells that compact during rearrangement in the pro–B-cell stage. Several transcription factors have been reported to be required for proper *Igh* locus compaction, including Pax5, YY1, and Ikaros (7, 11, 12). These proteins have also shown to be essential for the rearrangement of distal V_H J558 genes, but are not required for proximal V_H gene rearrangement. This process has led to the paradigm that the process of locus compaction is responsible for bringing the V_H genes in the middle and distal portions of the *Igh* locus into proximity to the single rearranged DJ_H gene that may be up to 2.4 Mb away from the incoming V_H gene. However, the precise manner in which these proteins result in locus compaction and distal V_H gene rearrangement is unknown. In addition to the role of these transcription factors in controlling distal V_H gene rearrangement and locus compaction, proteins involved in higher-order chromatin structure and nuclear architecture may also be involved. Recently we have shown that a ubiquitously expressed insulator protein CTCF was bound at many sites within the *Igh* locus, and we therefore hypothesized that CTCF would play a role in *Igh* locus 3D structure and contraction (13, 14). Indeed, knockdown of CTCF did result in a more extended *Igh* locus in pro-B cells (15).

The noncoding germ-line transcription from unrearranged V_H genes and across the J_H-Cμ regions of the *Igh* locus in pro-B cells was first reported over 25 y ago, and these observations gave rise to the accessibility hypothesis (3, 4). Transcription through the J_H-Cμ region begins at two sites: the μ⁰ transcript starts upstream of the most 3′ D_H gene, DQ52, and the Iμ transcript begins at the intronic enhancer Eμ (4, 16). Transcription in the V_H part of the locus has been detected in both the sense (same transcriptional orientation as the V_H genes) and, more recently, antisense

direction (3, 17). The V_H sense transcripts are easily seen at V_HJ558 genes, and sense germ-line transcripts from other V_H families are present, but are far more difficult to detect (3, 18). The sense transcription at V_H regions begins at the V_H gene promoter and ends shortly after the end of the coding region for the few transcripts that have been cloned, whereas antisense transcription is spread over intergenic regions (3, 17). The antisense transcription in the V_H locus initiates in pro-B cells after DJ_H recombination is accomplished (17). The timing of both sense and antisense germ-line transcription precisely correlates with the timing of rearrangement, and thus has been proposed to make the regions accessible for rearrangement (2, 3). The regulatory elements and transacting factors for the antisense transcripts are not known, other than the recently described Pax5-activated intergenic repeat (PAIR) elements (19). These 14 PAIR elements contain binding motifs for Pax5, CTCF, and E2A. Transcription from PAIR elements is dependent on the transcription factor Pax5, which only binds to these elements in pro-B cells (19). Data from our laboratory demonstrated that CTCF acts as negative regulator of transcription from PAIR elements because the knockdown of CTCF resulted in increased antisense transcription from PAIR elements in pro-B cells (15). However, the direct role of antisense transcription in V(D)J recombination has not been determined.

In the present study, using directional RNA-Seq analysis of the *Igh* locus, we show that the distal V_H portion of the *Igh* locus shows extensive antisense transcription, which is primarily localized downstream of 3 of the 14 PAIR elements. Importantly, we demonstrated that the PAIR promoter regions of antisense transcription in the distal V_H region directly interact with Eμ, the start site of the Iμ transcript, even though the PAIR transcription was Eμ-independent. These transcription-mediated DNA loops were made in a strict lineage- and stage-specific manner. Antisense transcription and PAIR-Eμ interactions were drastically reduced in $YY1^{-/-}$ pro-B cells, which are impaired in *Igh* locus compaction and distal V_H gene rearrangement (11). Because transcription takes place in discrete nuclear structures called transcription factories (20, 21), we hypothesize that localization in a common transcription factory may be the mechanism by which these distal regions came into proximity with Eμ, and this colocalization may initiate locus compaction. We propose that transcription-mediated long-range interactions at transcription factories is one of the key roles of noncoding antisense and sense transcription during V(D)J recombination, and that the transcription might initially alter the 3D conformation of the *Igh* locus, bringing the distal part of the locus close to Eμ, which is within ~2 kb of the DJ_H rearrangement to which one of the V_H genes will recombine.

Results

RNA-Seq Analysis of Sense and Antisense Transcription of *Igh* Locus.
Noncoding germ-line transcription from unrearranged V_H genes and intergenic regions of the *Igh* locus have been well documented (3, 17); however, no detailed genome-wide study has been done to quantitate and map all of the sense and antisense germ-line transcripts in the *Igh* locus. Therefore, we performed directional RNA-seq that can identify transcripts in both sense and antisense orientations using RNA from freshly isolated $Rag1^{-/-}$ pro-B cells. The number of reads from the first run clearly showed all of the major noncoding transcripts, but was not high enough to allow analysis of weak transcript levels (Fig. S1*A*). Therefore, we subsequently performed a pre-enrichment of RNA on custom Agilent arrays coated with all of the nonrepetitive DNA from the *Igh* locus, and ran a second RNA-seq experiment with this enriched RNA sample. This approach provided many more reads for the weaker V_H sense transcripts (Fig. S1*B*). Both RNA-seq analyses agreed in general. Surprisingly, by far the strongest transcription within the V_H portion of the locus was the antisense transcription in the V_HJ558 region, and most of these transcripts originated from PAIR4 and PAIR6 elements, with a smaller level from PAIR11 (Fig. 1). PAIR4 and PAIR6 also showed very strong Pol II binding and H3K4me3 (Fig. S1*D*). In addition, two lower

broad regions of antisense transcription were observed, one initiating sharply between V_H12.1.78 and J606.1.79, encompassing most of the J606 genes, and the other broad peak encompassing the more proximal V_HJ558 genes (Fig. S2). The latter is in the region where the original V_H antisense transcripts were described (17). We observed very strong peaks of sense transcripts through the μ⁰ and Iμ regions, as expected, and these were highly enriched in Pol II binding and H3K4me3 (4, 16) (Fig. S1*C*). We also observed low levels of sense transcripts throughout the V_H locus, the vast majority of which colocalized with V_H coding regions (Figs. S1 *A and B* and S3). Almost all of the individual functional V_H genes from the 7183, Q52, SM7, V_H10, and 3609 and families had some low level of sense transcripts, but many of the J558 genes and two of the SM7 genes, especially SM7.2.49, had somewhat higher levels of expression (Fig. S3). Although the sense transcripts were primarily confined to the coding region of V_H genes, there were some regions where sense transcripts were also found in intergenic regions, especially in the J558 region. Two intergenic regions, between 7183.7.10 and PG.1.11 and between J558.87.193 and J558.88.194, had relatively high levels of sense transcription (Fig. S3*B*).

Greatly Reduced Germ-Line Transcription in the Distal V_HJ558 Region in $YY1^{-/-}$ Pro-B Cells.
YY1 is an important regulator of B-cell development. $YY1^{-/-}$ pro-B cells have a block in differentiation, and have greatly impaired *Igh* locus compaction and distal V_HJ558 gene rearrangement (11). Therefore, we examined sense and antisense transcription in $YY1^{-/-}$ pro-B cells. As shown in Fig. 2*A*, transcription of PAIR4 and PAIR6 was reduced >95% in $YY1^{-/-}Rag1^{-/-}$ pro-B cells compared with $Rag1^{-/-}$ pro-B cells. Antisense transcription within the proximal half of the V_HJ558 region and in the J606 region were reduced >90% in $YY1^{-/-}Rag1^{-/-}$ pro-B cells. Sense transcription from the V_HJ558 region was also greatly reduced in $YY1^{-/-}Rag1^{-/-}$ pro-B cells (Fig. 2*A*). These results clearly demonstrate that YY1 acts as positive regulator of both sense and antisense transcription in the distal V_H region.

YY1 has been shown to bind to the Eμ enhancer (22). Because $YY1^{-/-}$ pro-B cells do not undergo *Igh* locus compaction and do not express sense or antisense transcripts initiating in the distal V_H portion of the locus, we wished to determine if YY1 bound to sites other than Eμ in the *Igh* locus, particularly in the V_H region. We therefore performed ChIP-seq for YY1 using freshly isolated pro-B cells from $Rag1^{-/-}$ mice (Fig. 2*B*). As expected, Eμ was the most enriched site in the *Igh* locus for YY1 binding, and the level of binding was low throughout the 2.4-Mb V_H portion of the *Igh* locus. To determine more precisely if there was truly a low level of YY1 binding within the V_H locus, we tested several locations, including small peaks identified in the ChIP-seq, by conventional ChIP/ quantitative PCR (qPCR) assay in $Rag1^{-/-}$ pro-B cells (Fig. 2*C*). Our results showed that 3′RR/CTCF (hs7, hs6) and DFL/CTCF regions had little enrichment of YY1 binding. However, some regions within the V_H locus showed low level enrichment of YY1. Four of the peaks identified in the ChIP-seq (V_HA, V_HB, V_HC, and V_HD) exhibited some low-level YY1 binding by ChIP/ qPCR. One of those sites, V_HB, was located at the start of the antisense transcription at the J606 region. PAIR elements also showed YY1 binding. Because antisense transcription in the distal V_HJ558 region was greatly reduced in $YY1^{-/-}$ pro-B cells, we examined the binding of YY1 on the previously described J558 5′ and 3′ intergenic regions but observed minimal enrichment by ChIP/qPCR at these sites. However, these primer sites are unlikely to be in the promoters. Other sites [CTCF binding sites V_H1, V_H3, V_H10 (13)] with no evidence of YY1 binding by ChIP-seq were tested and were negative. Thus, taken together, our results demonstrate that YY1 binds at low levels to several sites within the V_H locus, including near the start sites of some of the major antisense transcripts, and that either directly, by binding to these sites, or indirectly, YY1 is essential for noncoding transcription in the distal half of the V_H locus.

Because transcription from PAIR elements has been shown to require Pax5 binding (19), we tested Pax5 expression levels by
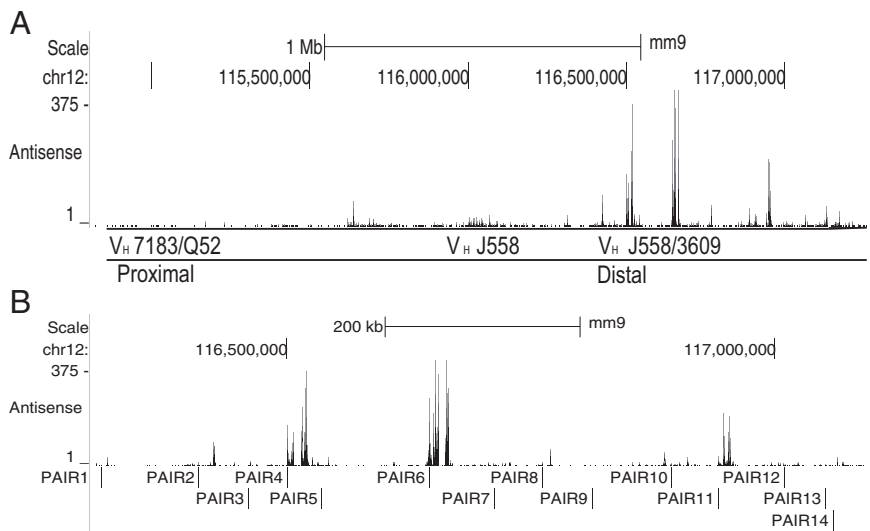
Fig. 1. RNA-seq analysis of the *Igh* locus. (*A*) RNA-seq was performed on *Rag1*$^{-/-}$ pro–B-cell RNA enriched on a custom Agilent array coated with the entire nonrepetitive DNA of the *Igh* locus. Antisense transcripts throughout the V$_H$ region of *Igh* locus are shown. (*B*) Mapping of antisense transcription in distal half of the V$_H$J558 region. The relative position of all 14 PAIR elements is shown.

qPCR (Fig. 2*D*) and demonstrated that there was a twofold reduction in the level of Pax5 mRNA in *YY1*$^{-/-}$*Rag1*$^{-/-}$ pro-B cells compared with both *Rag1*$^{-/-}$ pro-B cells (11). Next, we examined Pax5 binding by ChIP/qPCR and observed ~twofold reduction in Pax5 binding at PAIR4 and PAIR6 in *YY1*$^{-/-}$ pro-B cells compared with *Rag1*$^{-/-}$ pro-B cells. Binding was also reduced at known Pax5 sites in CD19 and BLNK (Fig. 2*E*). Whether this result is because of YY1 regulation of Pax5, or because of an earlier block in B-cell differentiation in *YY1*$^{-/-}$ pro-B cells compared with *Rag1*$^{-/-}$ pro-B cells is not known.

Long-Range Chromosomal Interactions of PAIR Elements with Eμ. RNA Pol II-mediated transcription occurs in discrete nuclear structures called transcription factories, and many genes are transcribed in each transcription factory (20, 21). Thus, if the Iμ transcription at Eμ and the antisense transcription are taking place in the same transcription factory, which would be reasonable because they are within 2 Mb of genomic distance, then the act of moving to the same transcription factory could directly result in looping and contraction of the *Igh* locus, bringing distal V$_H$ genes close to the DJ$_H$ rearrangement that is adjacent to Eμ. Therefore, our finding that the presence or absence of antisense transcription in the distal V$_H$ region is correlated with the extent of distal V$_H$ gene recombination led us to hypothesize that the transcription machinery may be actively involved in establishing long-range DNA interactions between promoters in the distal V$_H$ region and Eμ. To address this hypothesis, we used a quantitative chromosome conformation capture (3C) assay to ask if there were specific direct interactions between Eμ and the most highly transcribed regions in distal V$_H$ locus, PAIR4 and PAIR6 (Fig. 3*A*). Our analysis revealed interactions between both PAIR elements and Eμ in freshly isolated *Rag1*$^{-/-}$ pro-B cells (Fig. 3*B*). To determine whether or not these loops were present before the pro–B-cell stage, we also tested these interactions in *E2A*$^{-/-}$ pre-pro-B cells and found that these interactions were absent. In addition, thymocytes and murine embryonic fibroblasts also showed minimal PAIR–Eμ interactions (Fig. 3*B*). All of these other cell types do not make PAIR-directed transcripts. We conclude that interactions between PAIR elements and Eμ occur in a developmental stage- and lineage-specific manner, and correlate precisely with the transcriptional activity of the PAIR elements.

To further examine whether the formation of DNA loops between PAIR elements and Eμ correlates with antisense transcription, we analyzed 3C interactions in primary *YY1*$^{-/-}$ *Rag1*$^{-/-}$ pro-B cells, which have greatly reduced levels of antisense transcripts. We observed that the PAIR4–Eμ and PAIR6–Eμ 3C interactions were greatly reduced in *YY1*$^{-/-}$ *Rag1*$^{-/-}$ pro-B cells compared with *Rag1*$^{-/-}$ pro-B cells (Fig. 3*C*). However, the interaction between CTCF/DFL and CTCF/3′RR remained largely unaffected in *YY1*$^{-/-}$ *Rag1*$^{-/-}$ pro-B cells, as expected because these loops are CTCF-mediated (Fig. S4). Because CTCF knockdown leads to increased PAIR antisense transcription (15), we examined these PAIR–Eμ interactions in CTCF knocked-down pro-B cells. The 3C assay requires large numbers of cells in which CTCF has been knocked down, precluding the use of ex vivo pro-B cells. We therefore transduced a *Rag2*$^{-/-}$ pro–B-cell line with retroviruses expressing control or CTCF shRNA. The cells were harvested 4 d after transduction and prepared for 3C analysis. We observed that CTCF knock-down modestly enhanced 3C interactions between the PAIR elements and Eμ, although it did not achieve statistical significance (Fig. 3*D*), and it reduced the interaction between CTCF/DFL and CTCF/3′RR, as expected (Fig. S4). Thus, the extent of DNA looping between PAIR elements and Eμ in both CTCF knockdown pro-B cells and in *YY1*$^{-/-}$ pro-B cells correlates with the level of antisense transcription at the PAIR elements.

When promoters are dependent upon enhancers, those two regulatory elements often interact. Therefore, we wanted to determine whether the Eμ enhancer is required for the antisense transcription in the distal and proximal V$_H$J558 region. We assayed antisense transcription in pro-B cells from core Eμ$^{-/-}$ mice (23). Because these mice are on the 129 background, we compared them to control pro-B cells from *C.129 Rag1*$^{-/-}$ mice (Fig. 3*E*). These results clearly demonstrate that Eμ is not required for antisense transcription in distal V$_H$ region.

To more directly test the hypothesis that transcription, possibly at a transcription factory, is required for the long-range looping between Eμ and PAIR elements, we took advantage of published data demonstrating that blocking transcription initiation by heat shock can partially disrupt a transcription factory, whereas blocking transcriptional elongation will not disrupt an already existing interaction at a transcription factory (24). We therefore tested 3C interactions between PAIR elements and Eμ under these conditions. First, we treated primary *Rag1*$^{-/-}$ pro-B cells with heat shock. We confirmed that our heat-shock treatment worked by assaying the induction of the heat-shock–specific gene, HSP70 (Fig. 4*A*). As expected, the levels of PAIR antisense transcripts were reduced following heat shock. We demonstrated that the 3C interactions between the PAIR elements and Eμ were significantly reduced after the heat-shock treatment (Fig. 4*C*). Next, we blocked transcriptional elongation in pro-B cells with 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), which specifically blocks transcriptional elongation (25). We incubated pro-B cells with 100 μM DRB for 3 h, which greatly reduced PAIR transcription (Fig. 4*B*), and demonstrated that inhibition of transcription by DRB had no effect at all on long-
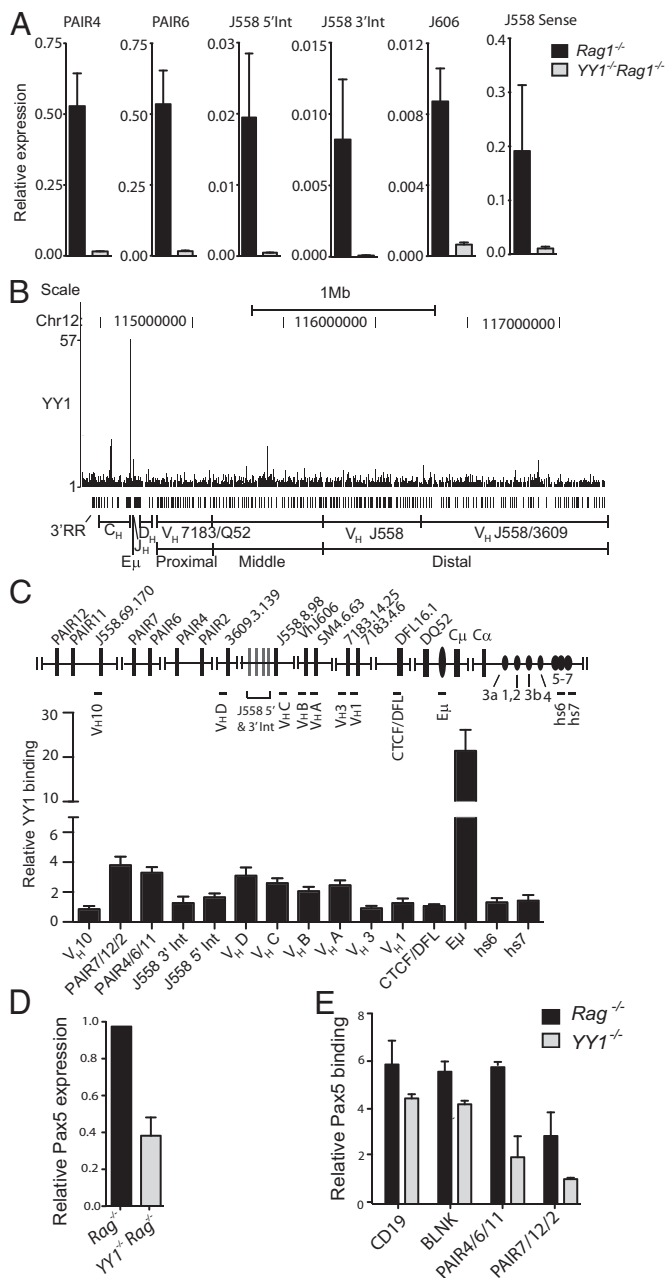
Fig. 2. Roles of YY1. (*A*) Antisense and sense transcription levels from RNA from freshly isolated pro-B cells from $Rag^{-/-}$ and $YY1^{-/-}Rag^{-/-}$ mice were measured by qPCR. Results are presented as the ratio of the amount of transcription in $YY1^{-/-}Rag^{-/-}$ pro-B cells compared with $Rag^{-/-}$ pro-B cells, and are the mean ± SEM ($n = 4$–5). (*B*) ChIP-seq analysis of YY1 binding throughout the *Igh* locus in $Rag1^{-/-}$ pro-B cells. Below the ChIP-seq Genome Browser picture, the vertical lines indicate the positions of V, D, and J genes. Note that the *Igh* orientation is the opposite from Fig. 2C. (*C*) ChIP/qPCR of indicated regions for YY1 binding in $Rag1^{-/-}$ pro-B cells. The schematic representation of the *Igh* locus showing the relative locations of regions tested in ChIP is presented in the panel directly above. Results are presented as mean ± SEM ($n = 5$). (*D*) RT-PCR analysis showing expression of Pax5 in $Rag^{-/-}$ and $YY1^{-/-}Rag^{-/-}$ pro-B cells. Results are presented as mean ± SEM from four to five independent experiments. (*E*) ChIP assay for enrichment of Pax5 at PAIR promoters. Results are presented as mean ± SEM from two independent experiments.

range DNA looping between PAIR elements and Eμ (Fig. 4D). This finding is consistent with the hypothesis that these interactions could be taking place in a transcription factory. This result also clearly reveals that antisense transcript itself is not

required for the maintenance of these long-range interactions, although its initial transcription is presumably required.

Discussion

Sense and antisense germ-line transcription at the *Igh* locus have been predicted to regulate V(D)J recombination by increasing the accessibility of *Igh* locus. Although it has been known for a long time that germ-line noncoding transcription takes place in pro-B cells, the extent and location of all of the sense and antisense transcription throughout the *Igh* locus has not previously been described. Here, using directional RNA-seq in $Rag1^{-/-}$ pro-B cells, we identified the full transcriptome of sense and antisense transcription throughout the *Igh* locus. In the sense direction, the $μ^0$ and Iμ transcripts were by far the most predominantly expressed, as expected. Surprisingly, there were only two major regions with a high level of antisense transcription, and they were located at the PAIR4 and PAIR6 elements in the distal $V_H3609/J558$ intergenic region. In addition, lower-level antisense transcription was observed at PAIR11, and even lower antisense transcription was seen in a broad region in the proximal half of the V_HJ558 region, and in the vicinity of the V_HJ606 genes. The level of sense transcripts at V_H genes was lower than that of the antisense transcripts. Sense transcription was higher at functional V_HJ558 and V_HSM7 genes than any other V_H genes in general, although most V_H7183, Q52, SM7, V_H10, and J558 functional
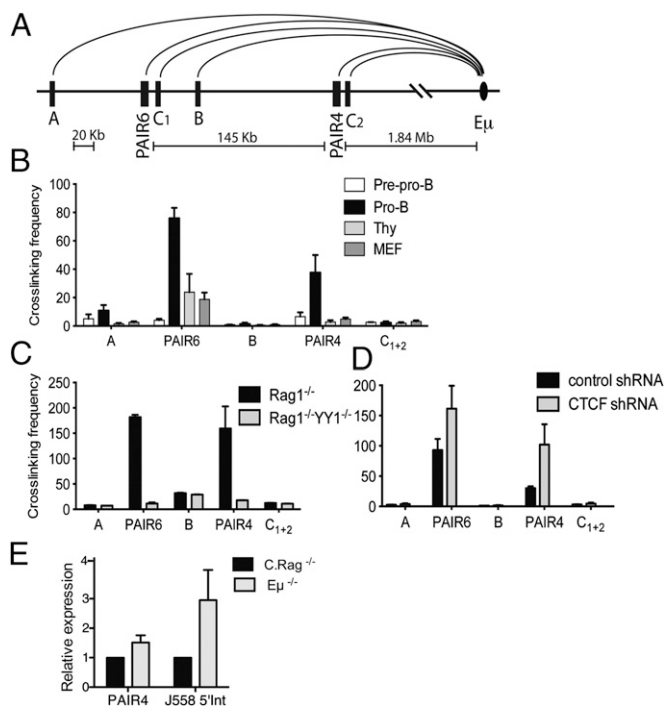
Fig. 3. The 3C analysis of long-range chromosomal loops between PAIR promoters and Eμ. (*A*) Schematic diagram of *Igh* locus showing the regions tested for 3C interactions and their relative distances. Primer pair C binds at two locations, indicated as C_1 and C_2. (*B*) The 3C analysis showing relative cross-linking frequencies between Eμ anchor fragment and HindIII fragments within the *Igh* locus using an Eμ TaqMan probe. Data are presented as mean ± SEM ($n = 2$–3). (*C*) The 3C analysis in $YY1^{-/-}$ $Rag1^{-/-}$ pro-B cells and $Rag1^{-/-}$ pro-B cells. Data are presented as mean ± SEM ($n = 3$). (*D*) The 3C analysis in $Rag2^{-/-}$ A-MuLV-transformed pro-B cells transduced with control or CTCF shRNA retroviruses. Data are presented as mean ± SEM ($n = 3$). (*E*) Antisense germ-line transcription is Eμ-independent. RNA was made from freshly isolated $CD19^+$ cells from $C.Rag1^{-/-}$ and $Eμ^{-/-}$ bone marrow, and were assayed for antisense transcription. The Igh^a locus of these 129 background mice does not amplify well with PAIR6, J558 3'Int primers, so only the PAIR4 and J558 5'Int data are displayed. Two RNA preparations of each genotype were assayed. Data are presented ± SEM.

IMMUNOLOGY
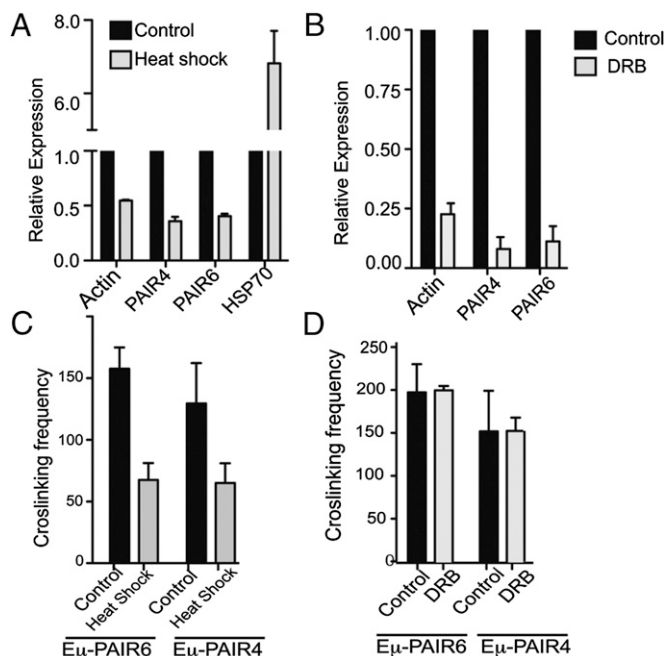
Fig. 4. Effect of transcription inhibition on long-range chromosomal loops in the *Igh* locus. *Rag1*$^{-/-}$ pro-B cells were subjected to heat shock at 45 °C for 30 min, or incubated with 100 μM DRB for 3 h. Cells were immediately harvested for RNA or for 3C lysate. (*A* and *B*) Relative expression levels following heat shock (*A*) and DRB (*B*) treatment on the indicated genes. Data are presented as mean ± SEM (*n* = 3). (*C* and *D*) The 3C analysis shows the relative cross-linking frequencies between the Eμ anchor fragment and PAIR6 and PAIR4. Data are presented as mean ± SEM (*n* = 4).

genes, but not pseudogenes, had some detectable sense transcription. Other V_H genes may have sense germ-line transcription, but at levels below the resolution of this RNA-seq.

Nascent transcription in the nucleus occurs in discrete sites, termed transcription factories, which contain high concentration of Pol II (20). Studies have shown that rather than the transcriptional machinery being recruited to and moving along the chromatin template, it is the genes that move to Pol II in the transcription factory (21). There is considerable amount of data suggesting that active coregulated genes from the same or different chromosomes often associate in the same factory (26–28). Given the precise timing of sense and antisense noncoding transcription, it seemed reasonable to propose that these transcriptionally active V_H regions, especially from the distal V_H3609/J558 part of the *Igh* locus, would be brought into close proximity to the highly transcribed Eμ region at a transcription factory, as previously suggested (29). Thus, transcription within transcription factories might be playing an important role in mediating the long-range chromatin interactions required to bring distal V_H genes in close proximity to the DJ_H locus during V_H to DJ_H recombination. Our results demonstrating DNA looping between PAIR elements and Eμ in pro-B cells support this hypothesis. The presence of these DNA loops only in pro-B cells, but not in pre-pro-B cells, thymocytes, or murine embryonic fibroblasts where noncoding transcription is absent, further suggested that these DNA loops are formed as a result of active transcription.

Additional support for this hypothesis came from other observations. First, we found that $YY1^{-/-}$ pro-B cells, which express negligible antisense transcripts and few sense transcripts, also demonstrated greatly reduced DNA looping between PAIR elements and Eμ regions. In addition, we observed reduced 3C interactions between PAIR elements and Eμ regions when pro-B cells were treated with a short heat shock. The heat-shock conditions used in our experiment have previously been demonstrated to partially dissociate genes from transcription factories

(24). These data are consistent with the hypothesis that transcription factory-mediated DNA looping occurs in the *Igh* locus, although this data should be interpreted with the caveat that the heat-shock treatment may affect other cellular functions. In contrast to the heat-shock treatment, it has previously been demonstrated that DRB treatment, which blocks transcription elongation, will not disrupt the interaction of DNA that is already in a transcription factory (24). DRB treatment showed a strong reduction in the level of antisense PAIR transcripts. However, this treatment did not affect the long-range DNA interactions between PAIR elements and Eμ, consistent with the hypothesis that these interactions are taking place in a transcription factory, although this is not the only interpretation of the data. The DRB result also demonstrates that the noncoding PAIR transcripts themselves do not mediate these long-range interactions. Together, the heat-shock treatment and DRB treatment data are consistent with the hypothesis that these PAIR–Eμ interactions take place in a transcription factory, but they do not constitute proof that this is the location where the interactions are occurring.

We suggest that once the strong antisense PAIR4 and PAIR6 regions are recruited into transcription factories along with Eμ, a subset of the many V_H sense promoters or other antisense promoters that are active in a given pro-B cell will also be recruited into the same transcription factory, bringing a variety of V_H genes in close proximity to Eμ region and its closely linked rearranged DJ_H region (Fig. S5). Hence, we propose that antisense and sense noncoding transcription plays an important role in the formation of the compacted recombination foci in which V(D)J recombination will subsequently take place. Our ChIP-seq analysis has shown that these transcriptionally active PAIR elements are also enriched for H3K4me3 modification (Fig. S1 *C and D*), which will aid in the recruitment of RAG2 (30, 31). Therefore, the transcription factory can also act as a platform to recruit the recombination machinery for V(D)J recombination. We predict that the transcription factory-dependent interactions of the many low-level sense-transcribed V_H genes and the weaker antisense transcribed regions will be stochastic, and that only a subset of V_H genes will be associated in a transcription factory in each pro-B cell at a given time. Assuming that the association of the V_H genes into transcription factories will be both random and dynamic, this would aid in the creation of a diverse repertoire of *Igh* rearrangements. This association could also account for the compacted 3D structure of the *Igh* locus during V_H-DJ_H rearrangement.

YY1 is known to be bound at Eμ (22), so the absence of YY1 binding to Eμ might be sufficient to result in the previously observed lack of locus contraction in $YY1^{-/-}$ pro-B cells (11). However, it has been shown that sense and antisense transcription in the distal V_H region does not depend upon Eμ (32, 33), and we showed here that PAIR antisense transcription is also independent of Eμ. Thus, we predicted that it is likely that YY1 has an additional, and perhaps a more direct, effect on the regulation of antisense transcription. We therefore performed ChIP-seq and ChIP/qPCR with YY1 on *Rag1*$^{-/-}$ pro-B cells. Indeed, ChIP/qPCR demonstrated a low level of enrichment at a few sites chosen from the ChIP-seq data, such as at the PAIR elements and near the promoter of the J606 antisense transcripts. Thus, we predict that one role of YY1 in inducing locus compaction may involve YY1 regulating antisense promoters, in addition to its role as a key Eμ binding transcription factor. It should be noted that our ChIP/qPCR results differ somewhat from the published results of Guo et al. (34). The difference may possibly be because of the fact that we analyzed freshly isolated *Rag1*$^{-/-}$ pro-B cells, whereas Guo et al. studied Abelson transformed cell lines. Furthermore, we observed that the level of Pax5 mRNA is twofold lower in $YY1^{-/-}$ pro-B cells, and binding of Pax5 to the PAIR promoters also reduced. Because Pax5 is essential for antisense transcription from PAIR elements (19), this could also contribute to the reduction in PAIR-supported antisense transcription. In addition, we demonstrated that $YY1^{-/-}$ pro-B cells showed greatly diminished interactions of PAIR elements with Eμ. We hypothesize that the absence of transcription

per se is what is responsible for the lack of PAIR–Eµ long-range interactions in $YY1^{-/-}$ pro-B cells, and hence lack of locus compaction.

Although we propose that the movement of sense and antisense promoters to the same transcription factory as Eµ would create a looped compacted structure, this is unlikely to be sufficient to endow full accessibility of V genes to subsequently undergo V(D)J rearrangement. The insertion of transcriptional terminator cassettes downstream of TEA or Jα56 showed that the rearrangement of the downstream 5–10 Jα genes was reduced, as was the level of AcH3 and H3K4me3 on those nontranscribed genes (35, 36). Thus, we propose a two-step model in which transcription facilitates the looping and contraction of the locus by bringing distal regions together with Eµ, possibly in a common transcription factory, in the first step (Fig. S5). This restructuring of the locus would be in the context of a locus that already has a CTCF-orchestrated 3D structure (15). This restructuring could be followed by additional stochastic and dynamic interactions from the many sense and antisense germ-line promoters throughout the *Igh* locus, further compacting the locus. Thus, this first step involves major changes in the 3D-structure of the locus, bringing the distal part of the locus close to Eµ (i.e., locus compaction) and we suggest it is facilitated by noncoding transcription as well as CTCF/cohesin-mediated loops. The second step would be the generation of accessibility at the recombination signal sequences (RSSs), which is likely to be accomplished by transcription through the V, D, or J genes and the RSS, which also endows the gene with H3K4me3 and thus will aid in RAG2 recruitment (30, 31). Furthermore, RAG1 binding is dependent upon intact promoters and enhancers as well as on transcription across the RSS (37).

Taken together, our experiments provide a direct test of the hypothesis that the transcription machinery itself, possibly via transcription factories, may act as a molecular tie, which plays a pivotal role in the first step of bringing the distal V_H region into close proximity to the DJ_H region to which one of the V_H genes will recombine. The presence of sense promoters at each functional V_H gene, and of many antisense promoters throughout the distal V_HJ558/3609 region provide a wealth of sites, a dynamic and stochastic subset of which could go to the Eµ-containing transcription factory in each pro-B cell, and thus position the neighboring V_H genes in an optimal location for rearrangement.

Materials and Methods

Mice, Cell Lines, and Pro–B-Cell Isolation. Mice and cell lines are described in *SI Materials and Methods*. $Rag^{-/-}$ and $YY1^{-/-}$ pro-B cells were obtained from bone marrow cells by isolation using CD19 Microbeads (Miltenyi). All pro-B cells were used directly ex vivo, except for the CTCF knockdown experiments.

RNA-Seq, ChIP, ChIP-Seq, and RT-PCR Analysis. For RNA-seq, 10 µg of total RNA was depleted of ribosomal RNA using RiboMinus (Invitrogen) kit. The details of fragmentation, linker ligation, and sequencing are provided in *SI Materials and Methods*. ChIP and ChIP-seq samples were prepared as previously described (15). The list of antibodies used is given in *SI Materials and Methods*, and primers in Tables S1 and S2.

CTCF Knockdown and 3C Analysis. The 3C analysis and the knockdown were performed as previously described (17). CTCF knockdown was performed on 4-d cultured pro-B cells with retroviral vectors containing CTCF shRNA target and control sequences. Additional details for both protocols are provided in the *SI Materials and Methods* and Table S3. See Table S4 for statistical analysis.

1. Alt FW, et al. (1984) Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J* 3:1209–1219.
2. Yancopoulos GD, Alt FW (1986) Regulation of the assembly and expression of variable-region genes. *Annu Rev Immunol* 4:339–368.
3. Yancopoulos GD, Alt FW (1985) Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. *Cell* 40:271–281.
4. Lennon GG, Perry RP (1985) C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5′-nontranslatable exon. *Nature* 318:475–478.
5. Johnston CM, Wood AL, Bolland DJ, Corcoran AE (2006) Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. *J Immunol* 176:4221–4234.
6. Kosak ST, et al. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296:158–162.
7. Fuxa M, et al. (2004) Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* 18:411–422.
8. Sayegh CE, Jhunjhunwala S, Riblet R, Murre C (2005) Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev* 19:322–327.
9. Jhunjhunwala S, et al. (2008) The 3D structure of the immunoglobulin heavy-chain locus: Implications for long-range genomic interactions. *Cell* 133:265–279.
10. Roldán E, et al. (2005) Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* 6:31–41.
11. Liu H, et al. (2007) Yin Yang 1 is a critical regulator of B-cell development. *Genes Dev* 21:1179–1189.
12. Reynaud D, et al. (2008) Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol* 9:927–936.
13. Degner SC, Wong TP, Jankevicius G, Feeney AJ (2009) Cutting edge: Developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* 182:44–48.
14. Degner-Leisso SC, Feeney AJ (2010) Epigenetic and 3-dimensional regulation of V(D)J rearrangement of immunoglobulin genes. *Semin Immunol* 22:346–352.
15. Degner SC, et al. (2011) 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 USA* 108:9566–9571.
16. Thompson A, Timmers E, Schuurman RK, Hendriks RW (1995) Immunoglobulin heavy chain germ-line JH-C mu transcription in human precursor B lymphocytes initiates in a unique region upstream of DQ52. *Eur J Immunol* 25:257–261.
17. Bolland DJ, et al. (2004) Antisense intergenic transcription in V(D)J recombination. *Nat Immunol* 5:630–637.
18. Love VA, Lugo G, Merz D, Feeney AJ (2000) Individual V(H) promoters vary in strength, but the frequency of rearrangement of those V(H) genes does not correlate with promoter strength nor enhancer-independence. *Mol Immunol* 37:29–39.
19. Ebert A, et al. (2011) 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* 34:175–187.
20. Sutherland H, Bickmore WA (2009) Transcription factories: Gene expression in unions? *Nat Rev Genet* 10:457–466.
21. Papantonis A, Cook PR (2011) Fixing the model for transcription: The DNA moves, not the polymerase. *Transcription* 2:41–44.
22. Park K, Atchison ML (1991) Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta), that binds to the immunoglobulin kappa 3′ enhancer and the immunoglobulin heavy-chain mu E1 site. *Proc Natl Acad Sci USA* 88:9804–9808.
23. Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E (2005) Elucidation of IgH intronic enhancer functions via germ-line deletion. *Proc Natl Acad Sci USA* 102:14362–14367.
24. Mitchell JA, Fraser P (2008) Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev* 22:20–25.
25. Marshall NF, Peng J, Xie Z, Price DH (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* 271:27176–27183.
26. Schoenfelder S, et al. (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* 42:53–61.
27. Brown JM, et al. (2006) Coregulated human globin genes are frequently in spatial proximity when active. *J Cell Biol* 172:177–187.
28. Xu M, Cook PR (2008) Similar active genes cluster in specialized transcription factories. *J Cell Biol* 181:615–623.
29. Matheson LS, Corcoran AE (2012) Local and global epigenetic regulation of V(D)J recombination. *Curr Top Microbiol Immunol* 356:65–89.
30. Matthews AG, et al. (2007) RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450:1106–1110.
31. Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S (2007) A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* 27:561–571.
32. Chakraborty T, et al. (2009) A 220-nucleotide deletion of the intronic enhancer reveals an epigenetic hierarchy in immunoglobulin heavy chain locus activation. *J Exp Med* 206:1019–1027.
33. Bolland DJ, et al. (2007) Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. *Mol Cell Biol* 27:5523–5533.
34. Guo C, et al. (2011) Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* 147:332–343.
35. Abarrategui I, Krangel MS (2006) Regulation of T cell receptor-α gene recombination by transcription. *Nat Immunol* 7:1109–1115.
36. Abarrategui I, Krangel MS (2007) Noncoding transcription controls downstream promoters to regulate T-cell receptor α recombination. *EMBO J* 26:4380–4390.
37. Ji Y, et al. (2010) Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination. *J Exp Med* 207:2809–2816.

IMMUNOLOGY