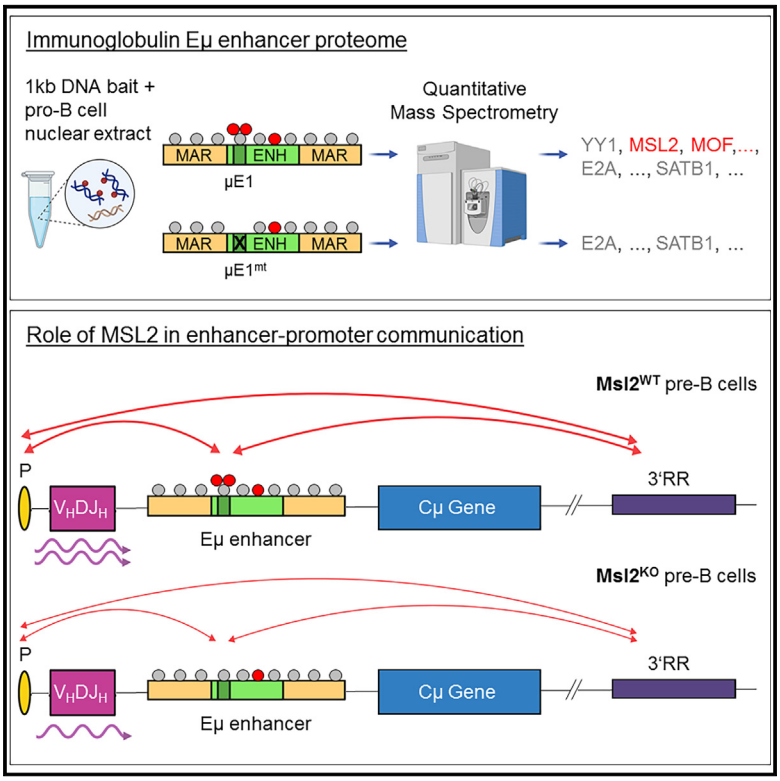
Cell Reports

Article

YY1-mediated enhancer-promoter communication in the immunoglobulin μ locus is regulated by MSL/MOF recruitment

Graphical abstract



Highlights

- Quantitative mass spectrometry of the immunoglobulin heavy chain Eμ enhanceosome
- The MSL/MOF complex is recruited to the Eμ enhancer via transcription factor YY1
- *Msl2* inactivation reduces μ expression and YY1-mediated enhancer-promoter looping
- *Mof* heterozygosity in mice impairs μ gene expression and early B cell differentiation

Authors

Yutthaphong Phongbunchoo, Fatima-Zohra Braikia, Cecilia Pessoa-Rodrigues, ..., Ranjan Sen, Gerhard Mittler, Rudolf Grosschedl

Correspondence

akhtar@ie-freiburg.mpg.de (A.A.), senranja@grc.nia.nih.gov (R.S.), mittler@ie-freiburg.mpg.de (G.M.), grosschedl@ie-freiburg.mpg.de (R.G.)

In brief

Phongbunchoo et al. perform quantitative mass spectrometric analysis of the immunoglobulin Eμ enhanceosome. They show that the identified MSL/MOF protein complex, a regulator of gene dosage compensation in *Drosophila*, binds Eμ via transcription factor YY1. MSL/MOF upregulates the expression of the functionally rearranged μ allele by augmenting YY1-mediated enhancer-promoter interactions.

Phongbunchoo et al., 2024, Cell Reports *43*, 114456
July 23, 2024
https://doi.org/10.1016/j.celrep.2024.114456

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

YY1-mediated enhancer-promoter communication in the immunoglobulin μ locus is regulated by MSL/MOF recruitment

Yutthaphong Phongbunchoo,[1] Fatima-Zohra Braikia,[2] Cecilia Pessoa-Rodrigues,[3] Senthilkumar Ramamoorthy,[1,4,5] Haribaskar Ramachandran,[1] Anna Grosschedl,[1] Fei Ma,[2] Pierre Cauchy,[1] Asifa Akhtar,[3,*] Ranjan Sen,[2,*] Gerhard Mittler,[6,*] and Rudolf Grosschedl[1,7,*]

[1]Laboratory of Cellular and Molecular Immunology, Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany
[2]Laboratory of Molecular Biology & Immunology, National Institute on Aging, NIH, Baltimore, MD, USA
[3]Department of Chromatin Regulation, Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany
[4]Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[5]Institute of Medical Bioinformatics and Systems Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[6]Proteomics Facility, Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany
[7]Lead contact
*Correspondence: akhtar@ie-freiburg.mpg.de (A.A.), senranja@grc.nia.nih.gov (R.S.), mittler@ie-freiburg.mpg.de (G.M.), grosschedl@ie-freiburg.mpg.de (R.G.)
https://doi.org/10.1016/j.celrep.2024.114456

SUMMARY

The rearrangement and expression of the immunoglobulin μ heavy chain (*Igh*) gene require communication of the intragenic Eμ and 3′ regulatory region (RR) enhancers with the variable (V_H) gene promoter. Eμ binding of the transcription factor YY1 has been implicated in enhancer-promoter communication, but the YY1 protein network remains obscure. By analyzing the comprehensive proteome of the 1-kb Eμ wild-type enhancer and that of Eμ lacking the YY1 binding site, we identified the male-specific lethal (MSL)/MOF complex as a component of the YY1 protein network. We found that MSL2 recruitment depends on YY1 and that gene knockout of *Msl2* in primary pre-B cells reduces μ gene expression and chromatin looping of Eμ to the 3′ RR enhancer and V_H promoter. Moreover, *Mof* heterozygosity in mice impaired μ expression and early B cell differentiation. Together, these data suggest that the MSL/MOF complex regulates *Igh* gene expression by augmenting YY1-mediated enhancer-promoter communication.

INTRODUCTION

Humoral immunity requires B lymphocytes to express single antigen-specific receptors, which is accomplished by the functional rearrangement of only one allele of immunoglobulin heavy chain (*Igh*) and light chain (*Igl*) genes. The large *Igh* locus is composed of many variable (V_H), multiple diversity (D), and four joining (J_H) gene segments and several constant (C) regions. In pro-B cells, the RAG-1 and -2 enzymes mediate the recombination of one of the D segments with one of the J_H segments upstream of the Cμ region by generating initially incompletely rearranged DJ_H-Cμ alleles. These alleles can be further recombined with one of the V_H gene segments to generate rearranged (V_HDJ_H)-Cμ alleles in pre-B cells.[1,2] The expression of a functionally rearranged μ allele results in the assembly of the pre-B cell receptor (pre-BCR), downregulation of the *Rag* genes, and the prevention of further rearrangement of the second *Igh* allele.[3] This process, termed allelic exclusion, ensures the monoallelic expression of antigen receptors.[4–6] In pre-B cells, the balance of signaling by the interleukin 7 receptor (IL-7R) and the pre-BCR regulates further rearrangement of the *Igl* loci and differentiation to immature B cells that express a functional BCR.[7–9]

The monoallelic expression of the functionally rearranged *Igh* allele is controlled by multiple layers of regulation. One of these involves the function of the tissue-specific intronic Eμ enhancer (Eμ), residing in the intronic region between the J_H gene segments and the Cμ region.[10,11] The Eμ enhancer, consisting of a 220-bp core enhancer and flanking nuclear matrix attachment regions (MARs), has been extensively studied and shown to confer chromatin accessibility, chromosome looping, gene rearrangement, and expression of the *Igh* locus.[12–20] Biochemical and functional studies using short Eμ core enhancer fragments identified several transcription factors (TFs) that bind multiple conserved enhancer sequence motifs, termed μE boxes; μA and μB motifs; and an Oct-binding site.[21] E-box-binding TFs include the basic-helix-loop-helix (bHLH) protein E2A (TCF3) and YY1, whereas the μA and Oct motifs are bound by ETS (Fli1 and Erg) and OCT proteins, respectively.[21] YY1 binds the μE1 box and has been shown to mediate 3D loop formation of Eμ with both the 3′ regulatory region (3′ RR) and rearranged V_H promoter regions.[18,22,23]

Experiments aimed at addressing the function of the Eμ enhancer in mice have shown that the deletion of the core enhancer in a functionally rearranged Ig μ transgene markedly impairs transcription from the rearranged V_H promoter.[12,24] Likewise, the deletion of the Eμ enhancer in the endogenous *Igh* locus reduces transcription of rearranged μ genes and impairs the efficiency of V_HDJ_H recombination by decreasing the accessibility of the D-J_H region.[15–17] Although the MARs flanking the Eμ enhancer core have been found to be dispensable for endogenous μ gene expression in germline mouse chimeras,[15] the MARs have been shown to potentiate the Eμ enhancer in activating the rearranged V_H promoter in transgenic mice and in pro-B cells that were transfected with *in vitro* CpG-methylated plasmid DNA.[13,25] Interestingly, a genomic footprinting analysis to assess Eμ enhancer occupancy in wild-type (WT) and ΔMAR transgenes revealed no detectable differences in mutant versus WT transgenic pro-B cells.[26] Thus, the question arises as to which Eμ-associated proteins are involved in the communication of the enhancer with the promoter of the rearranged μ gene.

To address this question, we investigated the protein network of the entire 1-kb Eμ enhancer region by combining state-of-the-art mass spectrometry with *in vitro* reverse chromatin immunoprecipitation (ChIP) pull-down of WT and mutant Eμ DNA fragments.[27,28] The analyzed datasets include novel DNA-binding proteins and higher-order protein complexes, among them components of the MSL/MOF multiprotein complex. In *Drosophila melanogaster*, the male-specific lethal (MSL) complex, consisting of the histone H4K16 acetyltransferase MOF, the MSL1–MSL3 proteins, MLE, and two non-coding RNAs, mediates the 2-fold upregulation of transcription from the single male X chromosome relative to that from two X chromosomes in females.[29,30] This function of the MSL/MOF complex compensates gene expression from the single male X chromosome relative to that from two X chromosomes in females. In addition, MOF is also a component of the non-specific lethal (NSL) complex, which includes KANSL1-3, MCRS2, and MBD-R2.[31,32] The NSL complex binds preferentially to promoters of constitutively active housekeeping genes, whereas the MSL complex binds to enhancers and gene bodies.[30,33–37] Notably, the mammalian MSL complex is not confined to the X chromosome but also binds to many developmentally regulated genes.[38] By cell-specific gene knockout of *Mof* in mice, this common component of the MSL and NSL complexes has been shown to be important for embryonic and neuronal stem cell and progenitor cell differentiation, hematopoiesis, and T cell differentiation.[39–41] Based on the well-documented role of MSL/MOF as a gene dosage regulator in *Drosophila*, we aimed to examine such a role in *Igh* "dosage compensation."

In the current study, we find that MSL2 and MOF are in complex with YY1 and are bound predominantly at the E1 box of the Eμ enhancer in pro-B and pre-B cells. Moreover, a *Msl2* knockout in pre-B cells or *Mof* heterozygosity in mice resulted in a 2-fold reduction in μ gene expression, reduction of active histone marks, reduced chromatin accessibility, and impaired RNA polymerase II (RNA POL II) recruitment. Interestingly, the *Msl2* knockout impaired Eμ enhancer-V_H promoter communication, suggesting that the MSL/MOF complex is a regulatory determinant of YY1-mediated chromatin looping.

RESULTS

Proteomic analysis of the immunoglobulin μ enhanceosome

To investigate the multiprotein complex bound at the intragenic Eμ enhancer, we adopted a quantitative mass spectrometry analysis of enhancer-binding proteins in pro-B cell nuclear extracts. To this end, we incubated a biotinylated 1-kb μWT DNA fragment, consisting of the 220-bp core enhancer, the Iμ promoter, and both 5′ and 3′ flanking MARs, with a nuclear extract of Abelson murine leukemia virus (A-MuLV)-transformed 38B9 pro-B cells. As a control for the sequence specificity of DNA-binding proteins, we used a DNA fragment in which the polarity of the DNA strands in their entire length (μREV) or in the core enhancer/Iμ promoter region (μCoreREV) had been inverted (Figure 1A). The inversion of the strand polarity was chosen to abrogate the binding by sequence-specific TFs without altering the interaction with repetitive elements as well as the AT/GC content of μWT. In six independent DNA pull-down experiments using forward and reverse SILAC (stable isotope labeling by amino acids in cell culture)[27] we reproducibly identified several known Eμ-binding TFs (TCF3, MEF2C, YY1 USF1, nuclear factor κB, FOXO1, ELF4, and IRF8) as well as previously unknown Eμ-binding proteins, including ZHX2, TFAP4, and CBFA2T2 (Figure S1A). Interestingly, we also identified multiple components of chromatin-regulatory complexes, including subunits (INO80, INO80c, and ACTR8) of the INO80 remodeler and the KANSL1 and KANSL3 subunits of the NSL/MOF complex (Figure S1A; Table S1).

To examine whether these proteins interact with the μ enhancer core or with flanking regions, we performed an additional and more sensitive mass spectrometry analysis of proteins. We assessed proteins differentially bound at the μWT bait versus the μCoreREV bait, containing the enhancer/Iμ promoter regions in an altered DNA sequence polarity. In five independent experiments, we detected an overlapping set of TFs and subunits of histone-modifying complexes that are bound specifically to the μWT bait, including the MSL2 subunit of the MSL/MOF complex (Figures 1B and S1B; Table S2).

To correlate the Eμ binding of proteins identified by the mass spectrometry analysis with enhancer function, we used Eμ enhancer mutations that had been shown to impair transcription of a rearranged μ gene.[24] Previously, we have shown that a deletion, removing part of the E1 box and the flanking 5′ MAR sequences (μΔE1ΔMAR), reduces transgene expression by almost three orders of magnitude relative to a WT transgene.[13,24] To assess the relative contribution of the μE1 box and the 5′ MAR, we generated additional transgenes carrying point mutations in the E1 box ($\mu E1^{mt1}$) or a deletion of the 5′ MAR (μΔMAR). Consistent with our previous analysis, the expression of the μΔE1ΔMAR transgene was markedly reduced at both the RNA and protein level (Figures 1C, 1D, and S1C). The $\mu E1^{mt1}$ and μΔMAR mutations reduced the expression of the μ transgene by 6- and 100-fold, respectively, suggesting a synergistic function of the μE1 box and MAR sequences.

We also assessed the effects of the mutations on the levels of active histone marks (H3K4me3 and H3K27ac) and inactive histone marks (H3K9me2, H3K9me3, and H3K27me3) at the V_H17.2.25 promoter of the rearranged μ transgenes and the μ

CellPress
OPEN ACCESS

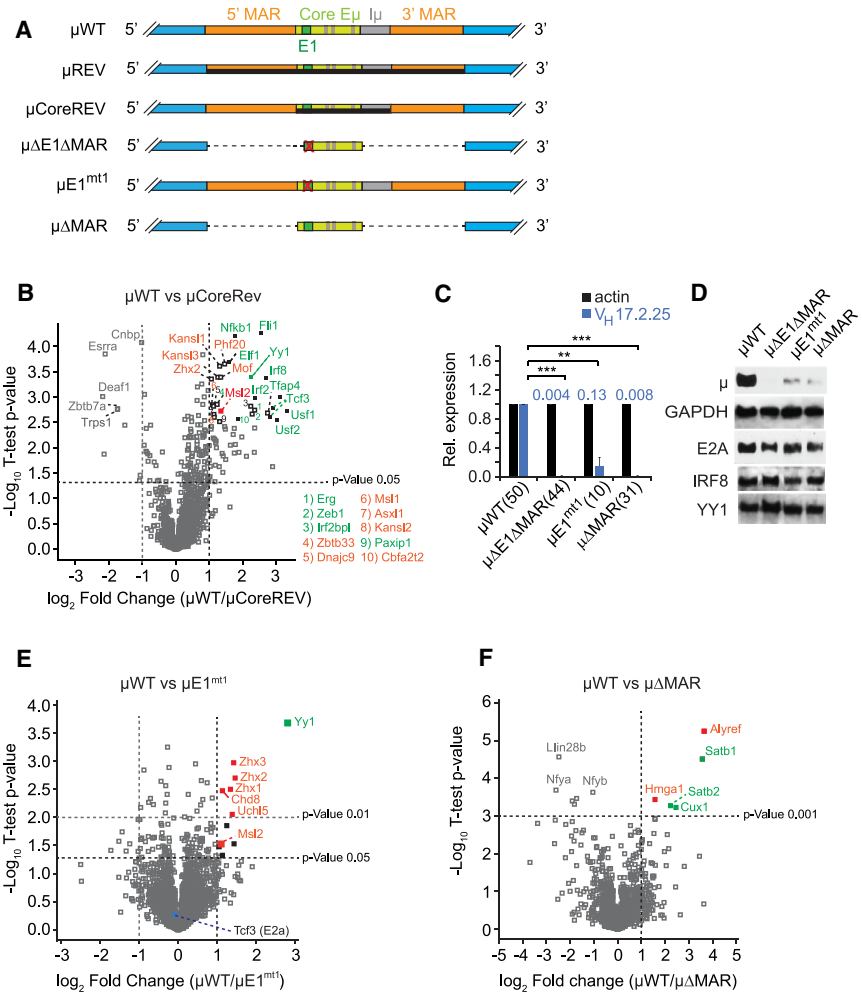


Figure 1. Determination of the Eμ enhancer proteome: μE1box- and μMAR-specific proteins

(A) Schematic of WT and Eμ enhancer mutant transgenes. The core enhancer, with the E1 box highlighted, is indicated in green, the Iμ promoter is shown in gray, and the matrix attachment regions (MARs) are indicated in orange. The reverse (REV) polarity of DNA segments is shown by a black bar. Deleted regions are shown by dashed lines.

(B) Volcano plot of proteins bound to the core Eμ enhancer. Green color represents known interaction partners. Red highlights novel candidate binders. Gray indicates proteins that preferentially bind to the mutant baits. Log_2 fold changes of μWT vs. μCoreREV <-1 or >1 and $p < 0.05$ are used as cutoff. The data represent five independent biological replicates.

(C) RT-qPCR analysis assessing μ mRNA expression from the V_H17.2.25 promoter of a rearranged WT or mutant μ transgene in A-MuLV-transformed fetal liver pro-B cells. The copy number of the transgene is shown in brackets. Numbers above bars indicate fold changes of the expression of mutant vs. WT transgenes. Error bars represent the standard deviation of three biological replicates. Statistical significance between WT and mutant cells was measured by an unpaired one-tailed Student's t test. $^{**}p < 0.01$, $^{***}p < 0.001$.

(D) Immunoblot analysis to detect μ protein expression in A-MuLV-transformed pro-B cells carrying a rearranged WT or mutant μ transgene. Expression of known Eμ enhancer-binding proteins and GAPDH served as a control. The blot is representative of three independent experiments.

(E) Volcano plot of proteins bound to the μE1 box of Eμ. Green color indicates known interaction partners. Red color indicates novel interaction partners. $p < 0.05$ was used as cutoff. The data represent six independent biological replicates.

(F) Volcano plot of proteins bound to μMARs of Eμ. Green color indicates known interaction partners. Red color represents novel μMAR candidate binding proteins. $p < 0.001$ cutoff applied. The data represent seven independent biological replicates.

enhancer core. In the μΔE1ΔMAR transgene, the active H3 histone marks K4me3 and K27ac were markedly reduced at the V_H promoter and modestly decreased at the Eμ core enhancer (Figure S2A). In this analysis, we included the promoter of the β-globin gene as a control for a non-expressed gene. The inactive histone mark H3K9me3 was found to increase predominantly at the Eμ core enhancer, suggesting that the enhancer core of the μΔE1ΔMAR transgene may acquire a bivalent chromatin state (Figure S2B). Consistent with the less severe effects on μ transgene expression, the individual μE1^{mt1} and μΔMAR mutations had modest effects on the levels of active and inactive histone marks (Figures S2A and S2B). We also examined the H4K16ac modification as an active mark that is deposited by the enzymatic activity of MOF and found modest but significant decreases of this histone mark at the enhancer and promoter of the mutant transgenes (Figure S2C).

To determine which proteins depend on the presence of the μE1 box or MAR sequences for Eμ enhancer binding, we performed an additional mass spectrometry analysis of DNA pull-down with μE1^{mt1} and μΔMAR DNA baits. By comparing the μWT versus μE1^{mt1} differential proteomes, we identified the known μE1-binding protein YY1 and, at lower abundance, members of the ZHX family of TFs, CHD8, and the MSL2 and MSL3 components of the MSL/MOF complex (Figures 1E and S2D; Table S3). In contrast to the WT vs. μCoreREV pull-down, the three KANSL members of the NSL/MOF complex were not differentially bound. Finally, we examined the differential association of proteins to the μWT versus μΔMAR baits and found MAR-specific binding of SATB1, SATB2, CUX, HMGA1, and ALYREF (Figures 1F and S2E; Table S4). Thus, we identified several new proteins that bind the μ enhancer in a μE1 box-dependent manner, raising the possibility that these proteins contribute to μE1-mediated regulation of immunoglobulin μ gene expression.

YY1 recruits the MSL2/MOF complex to the E1 box of the Eμ enhancer

To validate the binding of proteins identified by mass spectrometry, we repeated *in vitro* pull-down assays with μWT, μΔMAR,

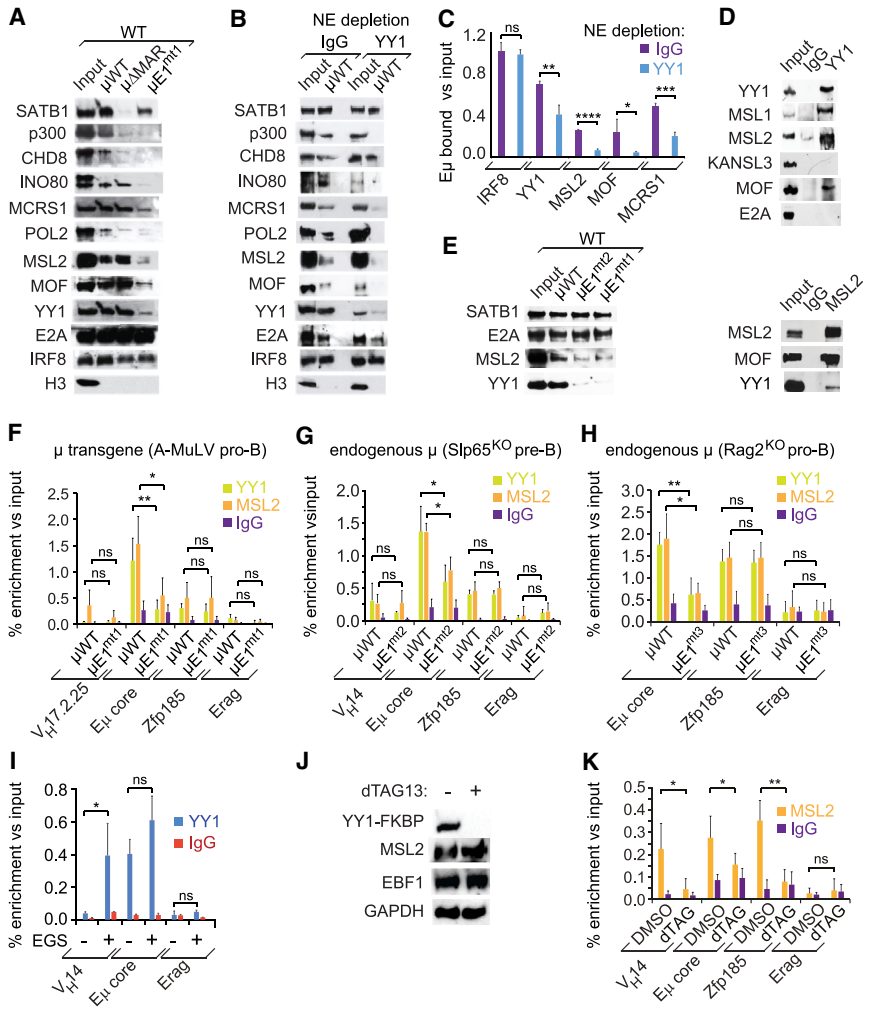


Figure 2. MOF/MSL complex proteins bind the μE1 box of the Eμ enhancer via YY1

(A) Immunoblot analysis of DNA pull-down with μWT, μΔMAR, or μE1^{mt1} DNA baits (Figure S1C) in 38B9 pro-B cell nuclear extracts to validate protein interactions as determined by mass spectrometry. The blot is representative of four independent experiments.

(B) Immunoblot analysis of DNA pull-down with μWT bait in WT or YY1-depleted 38B9 pro-B cell nuclear extracts (NE depletion). The blot is representative of three independent experiments.

(C) Quantification of the immunoblot (B) demonstrates YY1-dependent Eμ association of MSL2/MOF. IRF8 serves as control.

(D) Top: YY1 coIP with EGS-fixed *Slp65*^{KO} pre-B cell NE to assess interactions with components of the MSL and NLS complexes. The blot is representative of three independent experiments. Bottom: MSL2 coIP with EGS-fixed *Slp65*^{KO} pre-B cell NE to determine interactions with MOF and YY1. The blot is representative of three independent experiments.

(E) Immunoblot analysis of DNA pull-down with μWT, μE1^{mt2}, or μE1^{mt1} DNA baits (Figure S1C) in 38B9 pro-B cell NEs to assess whether the μE1^{mt2} mutation affects the binding of other proteins at Eμ (E2A) or MARs (SATB1). The blot is representative of three independent experiments.

(F) YY1 and MSL2 native ChIP in A-MuLV-transformed pro-B cell lines containing a rearranged μ transgene with μWT or μE1^{mt1} enhancers and a V_H17.2.25 promoter. *Zfp185* and *Erag* serve as positive and negative control, respectively. The values are normalized to the control loci.

(G) Native ChIP analysis to assess binding of YY1 and MSL2 at the Eμ enhancer and V_H14 promoter of the rearranged wild-type (WT) μ allele or the μE1^{mt2} mutant allele in *Slp65*^{KO} pre-B cells. In μE1^{mt2} cells, the E1 box of Eμ has been deleted by CRISPR-Cas9-mediated gene editing. Binding of both YY1 and MSL2 is reduced at the Eμ core in μE1^{mt2} cells relative to WT cells. *Zfp185* and *Erag* served as positive and negative control, respectively. The values are normalized to the control loci. The data represent three independent biological replicates.

(H) YY1 and MSL2 native ChIP in A-MuLV-transformed *Rag2*^{KO} cells containing WT or μE1^{mt3} *Igh* loci shows that YY1 and MSL2 bind to the μE1 box of the Eμ enhancer. The values are normalized to the control loci. Error bars represent the standard deviation of three biological replicates.

(I) Quantitative ChIP with or without EGS (bifunctional crosslinker) to detect YY1 binding at the V_H14 promoter or Eμ enhancer of the rearranged endogenous *Igh* locus in *Slp65*^{KO} pre-B cells. The IgG isotype serves as a negative control. The values are normalized to the control loci. The data represent three independent biological replicates.

(J) Immunoblot result showing successful degradation of YY1-FKBP after incubation with dTAG13 for 24 h. The blot is representative of four independent experiments.

(K) MSL2 ChIP in vehicle-treated (DMSO) and dTAG13-treated *Slp65*^{KO} pre-B cells carrying an FKBP-degron-tagged *Yy1* gene. The IgG isotype and *Actin* promoter serve as negative controls. The values are normalized to the IgG control. Error bars represent the standard deviation of three biological replicates. Statistical significance was determined by an unpaired one-tailed Student's t test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; ns, not significant.

and μE1^{mt1} baits and detected the proteins by immunoblot analysis. As expected, SATB1 binding was specifically impaired with the μΔMAR bait, whereas binding of INO80, YY1, MSL1, MSL2, and MOF was specifically reduced in the pull-down with the μE1^{mt1} bait (Figure 2A). Binding of p300 and CHD8 was diminished with both μΔMAR and μE1^{mt1} baits. YY1 binds directly μE1,[22] and therefore we examined whether YY1 mediates the recruitment of the MSL/MOF complex. To this end, we performed a DNA pull-down with nuclear extracts of pro-B cells in which YY1 had been partially depleted by addition of an anti-YY1 antibody (Figure 2B). In the pull-down with the YY1-depleted extract, we observed reduced binding of YY1, MSL2, MOF, MCRS1, p300, and RNA POL II relative to the pull-down with an anti-immunoglobulin G (IgG)-treated control extract (Figure 2B). Notably, the binding of IRF8 and SATB1 was not significantly changed by YY1 depletion (Figure 2B). The quantification of immunoblots confirmed the reduced binding of YY1, MSL2, MOF, and the MOF-associated protein MCRS1,[31] implying YY1-dependent recruitment of MSL2/MOF to Eμ (Figure 2C).

Therefore, we examined the interaction of YY1 with subunits of the MSL/MOF complex by co-immunoprecipitation (coIP). We detected an interaction of YY1 with MSL1, MSL2, and MOF

but not with KANSL3, suggesting that YY1 may preferentially recruit the MSL/MOF complex (Figure 2D, top). As a control, we did not detect an interaction of YY1 with E2A. Although an interaction of YY1 and MSL2 was also observed at a low level in a reciprocal coIP with anti-MSL2 (Figure 2D, bottom), these coIPs preclude any conclusion about direct or indirect interactions.

MSL2 and MOF were the most consistently enriched components of the MSL/MOF complex in the μWT vs. μE1^{mt1} mass spectrometry analysis. To obtain evidence for the recruitment of MSL2 to Eμ via YY1, we examined the effects of μE1 mutations and YY1 depletion on Eμ binding by MSL2. For the analysis of YY1 and MSL2 binding at the WT and μE1 box-mutated enhancer by ChIP analyses, we used three cellular models that offered distinct advantages. A-MuLV-transformed μ transgenic pro-B cells provided high copy numbers of a transgene. Non-transformed primary *Slp65*^{KO} pre-B cells and A-MuLV-transformed *Rag2*^{KO} pro-B cells enabled an analysis of protein binding at an endogenous rearranged μ gene and at μ germline alleles, respectively. *Slp65*^{KO} pre-B cells can be cultured in the presence of IL-7, and they express a functionally rearranged μ allele with a V_H14 gene segment.[42,43] In these three cellular models, different experimental strategies were used to generate the μE1 box mutation (see supplemental information for details), which resulted in slightly different but functionally equivalent mutations (μE1^{mt1}, μE1^{mt2}, and μE1^{mt3}; Figure S1C). In particular, *in vitro* pull-down experiments showed a similar reduction of YY1 and MSL2 binding to the μE1^{mt1} and μE1^{mt2} baits relative to the μE1^{WT} bait, whereas no changes in binding were observed for the MAR-binding protein SATB1 (Figure 2E).

In all three cell models, quantitative YY1 and MSL2 ChIP analysis indicated that binding of both proteins is significantly reduced at the Eμ core containing mutant μE1 sequences relative to the WT Eμ (Figures 2F–2H). At the V_H17.2.25 promoter of the μWT and μE1^{mt1} transgenes, we observed weak MSL2 binding, and at the V_H14 promoter of the rearranged μ gene, we detected weak YY1 and MSL2 binding, which was not significantly changed by the μE1^{mt2} mutation (Figures 2F and 2G). In these ChIP experiments, the *Erag* enhancer served as a negative control for YY1 and MSL2 binding. Likewise, the *Zfp185* locus served as a *bona fide* MSL2 control target[44] and YY1-bound gene. The V_H17.2.25 and V_H14 promoter regions lack discernable YY1 binding sites (data not shown), raising the possibility of binding site-independent DNA contacts of YY1 via looping to the μE enhancer. Therefore, we examined the effects of using a bifunctional crosslinker in ChIP analysis. By employing the EGS crosslinker together with formaldehyde, we observed a marked increase in the YY1 ChIP signal at the V_H14 promoter in *Slp65*^{KO} pre-B cells and a much weaker effect on YY1 binding at the Eμ core enhancer (Figure 2I). Thus, the association of YY1 with the V_H14 promoter is likely indirect via YY1-mediated looping to the enhancer.

We also examined the effects of YY1 depletion on MSL2 binding at Eμ. We could not adopt a YY1 knockout strategy in pro-B cells because the conditional inactivation of the *Yy1* gene in the B lymphoid lineage results in an early developmental block and precludes the establishment of YY1-deficient pro-B cell cultures.[45] Therefore, we generated *Slp65*^{KO} pre-B cells in which the FKBP degron was fused to the endogenous *Yy1* gene by CRISPR-Cas9 gene editing, allowing the inducible degradation of YY1 protein.[46,47] In cells that were treated with the heterobifunctional degradation TAG (dTAG) dimerizer, the YY1 protein was rapidly degraded after 24 h (Figure 2J). By quantitative MSL2 ChIP analysis with EGS-crosslinked control and dTAG-treated cells, we detected markedly reduced MSL2 binding at the Eμ enhancer and V_H14 promoter of the endogenous μ gene in dTAG-treated cells relative to DMSO control-treated cells. In the dTAG-treated cells, we also observed a marked reduction of MSL2 binding at the *Zfp185* locus, whereas no binding was observed at the *Erag* enhancer (Figure 2K). Thus, the depletion of YY1 both *in vitro* and *in vivo* leads to hampered binding of MSL2 at the Eμ core and mimics the effects of the μE1 box mutation. Together, these data indicate that the recruitment of MSL2 to the Eμ enhancer occurs primarily via YY1 binding at the μE1 box.

MSL2 and YY1 regulate μ gene expression

To examine whether MSL2 regulates μ gene expression, we generated *Msl2*^{KO} alleles in μ transgenic pro-B cells and in *Slp65*^{KO} pre-B cells by CRISPR-Cas9-mediated gene editing and by inserting a stop codon immediately downstream of the start codon. In μ transgenic pro-B cells, the absence of MSL2 resulted in markedly reduced μ expression, as assessed by immunoblot analysis and flow cytometry (Figures 3A and 3B). In the immunoblot analysis, we also found a reduction of the μ-associated VpreB surrogate light chain but no altered expression of YY1, KANSL3, or GAPDH. Notably, the re-expression of ectopic MSL2 resulted in an almost complete rescue of IgM expression, suggesting a direct role of MSL2 in μ gene expression (Figures 3A and 3B). By RT-qPCR analysis, we observed, in *Msl2*^{KO} pro-B cells, a, ∼10-fold decrease of transgenic μ RNA derived from the V_H17.2.25 promoter and a, ∼4-fold decrease of Eμ/Iμ-derived transcripts relative to *Msl2*^{WT} pro-B cells (Figures 3C and S3A). Consistent with the reduced μ transgene expression in *Msl2*^{KO} pro-B cells, we also detected modest but significant decreases in chromatin accessibility at the V_H17.2.25 promoter and Eμ enhancer (Figure 3D). The re-expression of MSL2 in *Msl2*^{KO} pro-B cells increased chromatin accessibility at the V_H17.2.25 promoter but not at the Eμ enhancer.

We further analyzed the effects of the *Msl2*^{KO} mutation on the expression of the endogenous rearranged μ allele in *Slp65*^{KO}*Msl2*^{KO} pre-B cells and observed a 2- to 4-fold decrease in the expression of endogenous μ protein relative to corresponding *Slp65*^{KO}*Msl2*^{WT} cells (Figure 3E). Flow cytometry analysis indicated that the decrease in intracellular μ protein expression is homogeneous in the *Msl2*^{KO} pre-B cell population (Figure 3F). Moreover, we noticed a reduced abundance of V_H14-specific μ transcripts and diminished chromatin accessibility at the rearranged V_H14 promoter (Figures 3G and 3H; Table S5). To determine whether the function of the μE1 box involves only the YY1-mediated recruitment of the MSL/MOF complex, we included in our analysis μE1^{mt2} pre-B cells. In these cells, μ protein and RNA expression was reduced more than 10-fold relative to μE1^{WT} cells (Figures 3E–3G; Table S5). The enhanced defect of the μE1 box mutation relative to the *Msl2* mutation may reflect an YY1 function by some other associated factors or an additional YY1-independent function of the μE1 box.

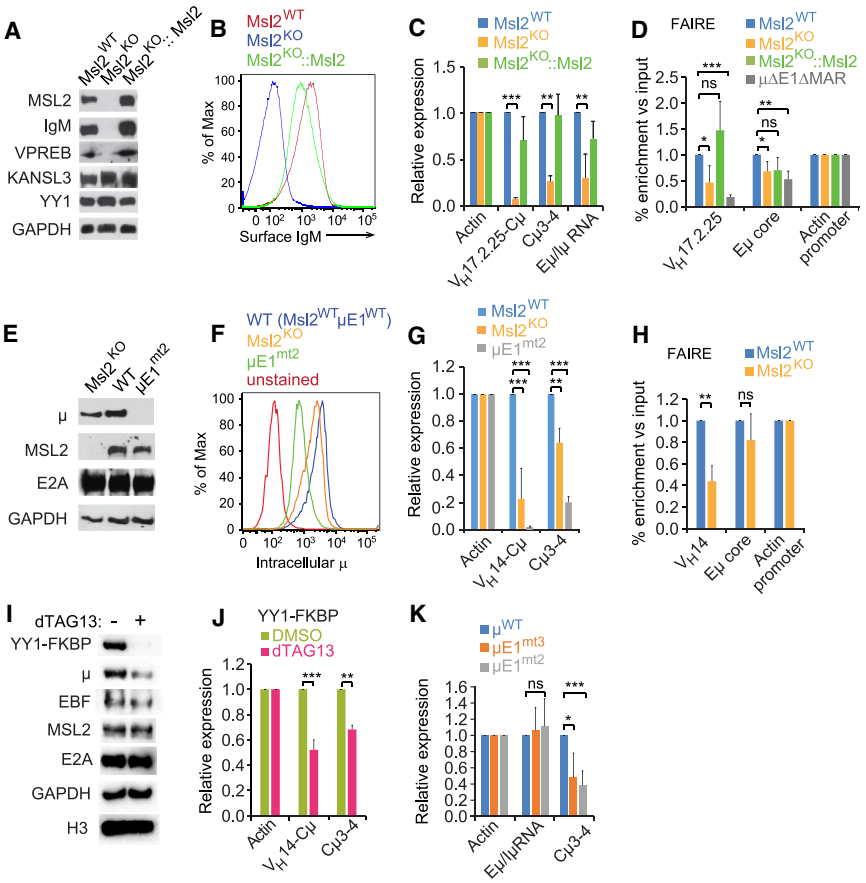


Figure 3. MSL2 binding of Eμ regulates transgenic and endogenous μ gene expression

(A) Immunoblot analysis to determine μ and surrogate light chain (VPREB) protein expression in μWT transgenic pro-B cells in which *Msl2* has been knocked out by CRISPR-Cas9-mediated gene editing. In *Msl2*^{KO} cells, μ expression is rescued by retroviral overexpression of ectopic MSL2-GFP (*Msl2*^{KO}::*Msl2*). The blot is representative of four independent experiments.

(B) Histogram of flow cytometry analysis of IgM expression on the surface of *Msl2*^{WT}, *Msl2*^{KO}, and *Msl2*^{KO}::*Msl2* pro-B cells carrying a μWT transgene. The data represent two independent biological replicates.

(C) RT-qPCR analysis of μWT transgene expression in *Msl2*^{WT}, *Msl2*^{KO}, and *Msl2*^{KO}::*Msl2* pro-B cells by using primers detecting transcripts from the rearranged V_H17.2.25 promoter (V_H17.2.25-Cμ), Cμ constant region (Cμ3-4), and Eμ-Iμ promoter (Eμ/Iμ RNA) region. The data represent five independent biological replicates.

(D) Analysis of chromatin accessibility of the V_H17.2.25 promoter and Eμ enhancer in *Msl2*^{WT}, *Msl2*^{KO}, and *Msl2*^{KO}::*Msl2* pro-B cells carrying a μWT transgene and in μΔE1ΔMAR μ transgenic pro-B cells by FAIRE (formaldehyde-assisted isolation of regulatory elements). The data represent five independent biological replicates.

(E) Immunoblot analysis to determine μ protein expression in primary *Slp65*^{KO} pre-B cells in which CRISPR-Cas9-mediated gene editing was used to knock out the *Msl2* gene or delete the μE1 box (μE1^{mt2}) of the endogenous rearranged μ allele. The blot is representative of three independent experiments.

(F) Histogram of flow cytometry analysis of intracellular μ protein in *Slp65*^{KO} pre-B cells containing an endogenous rearranged WT or μE1^{mt2} μ allele or *Msl2*^{KO} alleles. The data represent three independent biological replicates.

(G) RT-qPCR analysis of the rearranged μ gene with primers to detect V_H14-Cμ and spliced Cμ (Cμ3-4) transcripts. The data represent three independent biological replicates.

(H) FAIRE chromatin accessibility analysis of the rearranged V_H14 promoter and Eμ enhancer in *Slp65*^{KO} pre-B cell lines containing *Msl2*^{WT} or *Msl2*^{KO} genes. The data represent three independent biological replicates.

(I) Immunoblot analysis to detect expression of μ, MSL2, and YY1 in *Slp65*^{KO} pre-B cells in which the FKBP degron tag has been inserted into the endogenous *Yy1* gene by CRISPR-Cas9-mediated gene editing. YY1 protein degradation was induced by treatment of cells with dTAG13 for 24 h. The blot is representative of three independent experiments.

(J) RT-qPCR analysis of μ gene RNA expression in vehicle-treated (DMSO) or dTAG13-treated *Slp65*^{KO} pre-B cells carrying an *Yy1* gene with an FKBP degron tag. The data represent three independent biological replicates.

(K) RT-qPCR analysis to compare the effects of μE1^{mt2} and μE1^{mt3} mutations on the expression of the unrearranged μ gene in *Rag2*^{KO} pro-B cells. The data represent three independent biological replicates. Error bars represent the standard deviation of biological replicates. Statistical significance between WT and mutant cells was measured by an unpaired one-tailed Student's t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Therefore, we assessed the effects of induced YY1 degradation on the expression of the functionally rearranged μ allele in *Slp65*^{KO} pre-B cells. After 24 h of dTAG addition and degradation of YY1, we observed an approximately 2-fold decrease in the abundance of μ protein and V_H14-Cμ transcripts (Figures 3I and 3J). Thus, the lack of MSL2 recruitment to the *Igh* locus and the degradation of YY1 both result in a similar reduction in μ gene transcription. Despite the experimental differences of continuous *Msl2* inactivation versus short-term YY1 degradation, both conditions led to a less severe decrease in μ gene expression than the μE1 box mutation (Figures 3E, 3G, 3I, and 3J), suggesting that additional proteins act via the E1 box to regulate μ gene expression (see Figure S2D; Table S3 for potential candidates).

We also extended this analysis to *Rag2*^{KO} pro-B cells carrying unrearranged μ alleles. In comparison to cells with the μWT alleles, the μE1^{mt3} mutation resulted in an approximate 40% decrease in the abundance of spliced μ transcripts, detected with Cμ exon 3- and exon 4-spanning primers, but no difference in unspliced μ transcripts detected with intronic Eμ-Iμ primers (Figures 3K and S3A). We further introduced the μE1^{mt2} mutation into the unrearranged μWT gene of *Rag2*^{KO} pro-B cells and found that the μE1^{mt2} mutation had a similar effect on unrearranged μ gene expression as the μE1^{mt3} mutations (Figure 3K). In this experiment, we observed a differential effect of the μE1 mutation on Eμ-Iμ transcripts versus Cμ transcripts. This difference may be related to the splicing status of Cμ transcripts

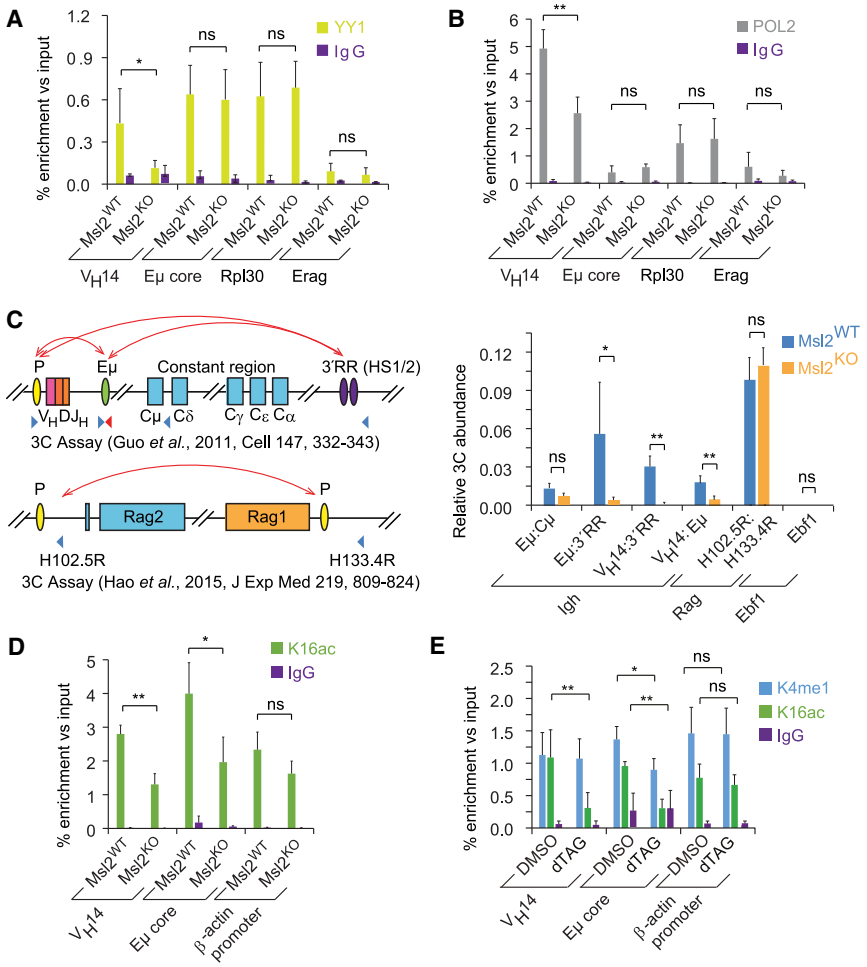


Figure 4. *Msl2* inactivation impairs YY1 interaction with the V_H promoter via DNA looping

(A) Quantitative ChIP with EGS fixation to detect YY1 binding at the V_H14 promoter or Eμ enhancer of the rearranged endogenous *Igh* locus in *Slp65*^{KO} pre-B cells containing *Msl2*^{KO} or *Msl2*^{WT} genes. IgG isotype serves as a negative control. *Rpl30* and *Erag* serve as positive and negative control loci, respectively. The data represent three independent biological replicates.

(B) RNA POL II ChIP in *Slp65*^{KO} pre-B cells containing *Msl2*^{KO} or *Msl2*^{WT} genes as described in (A). The data represent three independent biological replicates.

(C) Left: schematic of the rearranged *Igh* locus (top) and *Rag* locus (bottom) studied by 3C assays (right). The V_H14 promoter and Eμ are indicated by yellow and green ellipses, respectively. The *Igh* constant regions are shown as blue rectangles. The 3′ RR downstream enhancer (hypersensitive sites HS1/2) is highlighted in purple. P, promoter; E, enhancer; RR, regulatory region. Previously shown interactions are indicated by red arrows.[18,49] Primers used in the 3C assay are indicated by blue triangles. A short amplicon covering Eμ (utilizing the reverse primer depicted in red) served as normalization control. Right: 3C assays to assess MSL2-dependent interactions of the Eμ enhancer with the 3′ RR enhancer region (Eμ:3′ RR) and the rearranged V_H14 promoter (V_H14:Eμ) of the *Igh* locus in *Msl2*^{WT} and *Mls2*^{KO} *Slp65*^{KO} pre-B cells. Interactions of HindIII DNA fragments were analyzed by 3C, followed by qPCR detection. (V_H14:3′ RR) indicates interactions of the V_H14 promoter with the 3′ RR. As a control, no significant interactions between Eμ and Cμ were detected (Eμ:Cμ). 3C assays of the *Rag* and *Ebf1* loci served as negative controls (not influenced by *Msl2*^{KO}). The data represent three independent biological replicates.

(D) H4K16ac ChIP in *Slp65*^{KO} pre-B cells containing *Msl2*^{KO} or *Msl2*^{WT} genes. The IgG isotype served as a negative control. The data represent three independent biological replicates.

(E) H3K4me1 and H4K16ac ChIP in vehicle-treated (DMSO) or dTag13-treated *Slp65*^{KO} pre-B cells containing an FKBP degron-tagged *Yy1* gene. The β-actin promoter served as negative control. The values are normalized to the IgG control. Error bars represent the standard deviation of biological replicates. Statistical significance between WT and mutant cells was measured by an unpaired one-tailed Student's t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

versus Eμ-Iμ transcripts. To examine whether the μE1^{mt3} mutation causes an altered repertoire of rearranged V_H segments, we isolated pro-B cells from *Rag2*^{WT}μ^{WT} and *Rag2*^{WT}μE1^{mt3} mice and determined the relative V_H usage by deep sequencing.[48] This analysis indicated that the use of V_H gene families is similar in WT and μE1 mutant mice (Figures S3B and S3C).

MSL2 is involved in YY1-mediated chromatin looping of the μ enhancer

YY1 can act as a pioneering factor and regulates enhancer-promoter looping.[47] Quantitative YY1 ChIP analysis revealed similar YY1 binding at the endogenous Eμ core enhancer in *Msl2*^{KO} and *Msl2*^{WT} pre-B cells, suggesting that, by and large, MSL2 does not contribute to the E1 box-dependent binding of YY1 (Figure 4A). In this experiment, we used the EGS bifunctional cross-linker to examine a potential effect of MSL2 on YY1 binding at the V_H14 promoter. We observed a decrease of YY1 binding at the V_H14 promoter in *Msl2*^{KO} cells relative to *Msl2*^{WT} cells, whereas no significant differences were detected at the *Rpl30*-positive and *Erag*-negative controls. We also examined the effects of the *Msl2* knockout on the recruitment of RNA POL II and found reduced RNA POL II binding at the V_H14 promoter but not at the Eμ core enhancer (Figure 4B). Similarly, RNA POL II binding at the $V_H17.2.25$ promoter of the μWT transgene in *Msl2*^{KO} pro-B cells and at the $V_H17.2.25$ promoter of the μE1^{mt1} transgene in *Msl2*^{WT} pro-B cells was decreased relative to the WT controls (Figures S4A and S4B). Thus, the association of YY1 with the V_H14 promoter may occur via MSL2-enhanced chromatin looping by μE1-bound YY1.

To examine a potential role of MSL2 in chromatin looping, we performed a 3C assay in *Slp65*^{KO} pre-B cells, using anchors and primers as described previously.[18] V_H promoter activation by Eμ is also dependent on interactions of Eμ with the hypersensitive sites HS1/2 of the downstream 3′ RR.[18] By assessing the easily detectable long-range chromatin loops between Eμ and 3′ RR (HS1/2), we found that these long-range enhancer-enhancer interactions are significantly reduced in *Msl2*^{KO} cells relative

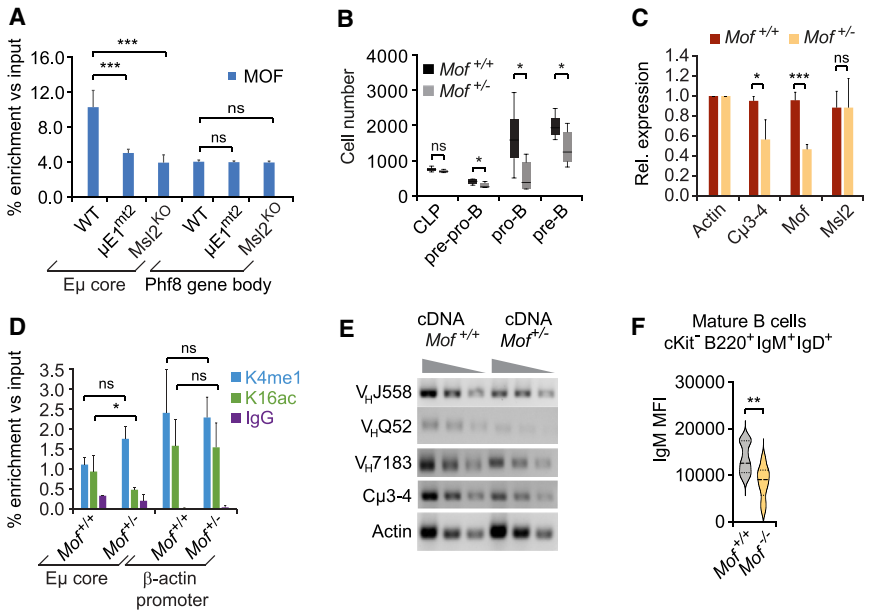


Figure 5. MOF regulates μ gene expression in mice

(A) MOF ChIP analysis in *Slp65*^{KO} pre-B cells containing a rearranged WT or μE1^{mt2} μm allele or *Msl2*^{KO} alleles reveals MOF association at the μE1 box of the Eμ enhancer. MOF binding at Eμ is impaired in *Msl2*^{KO} cells. The *Phf8* gene body serves as a control. The data represent four independent biological replicates.

(B) Flow cytometry analysis of the total numbers of common lymphoid progenitors (CLPs) and early B-lineage cells derived from *Mof* $^{+/+}$ or *Mof* $^{+/-}$ mice. MFI, median fluorescence intensity. The data represent three biological replicates.

(C) RT-qPCR analysis to determine mRNA expression of μ, *Mof*, and *Msl2* in pro-B cells derived from *Mof* $^{+/+}$ or *Mof* $^{+/-}$ mice, sorted by fluorescence-activated cell sorting and cultured for 3 days with IL-7 on OP-9 feeder cells. The data represent three biological replicates.

(D) ChIP analysis of active histone marks (H3K4me1 and H4K16ac) at the Eμ enhancer in A-MuLV-transformed *Mof* $^{+/+}$ or *Mof* $^{+/-}$ pro-B cells. The actin promoter serves as a negative control. The values are normalized to the control loci. The data represent three biological replicates.

(E) Semi-quantitative RT-PCR analysis to assess transcription from rearranged μ genes containing different V_H gene segments as determined by different V_H promoter-Cμ primers and visualized by agarose gel electrophoresis. The gel is representative of two independent experiments.

(F) Flow cytometric analysis of IgM expression in sorted cKit$^-$B220$^+$IgM^{lo}IgD^{hi} mature B cells derived from *Mof* $^{+/+}$ or *Mof*$^{+/-}$ mice. MFIs from two independent experiments are shown. Error bars represent the standard deviation of biological replicates. Statistical significance between WT and mutant cells was measured by an unpaired one-tailed Student's t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

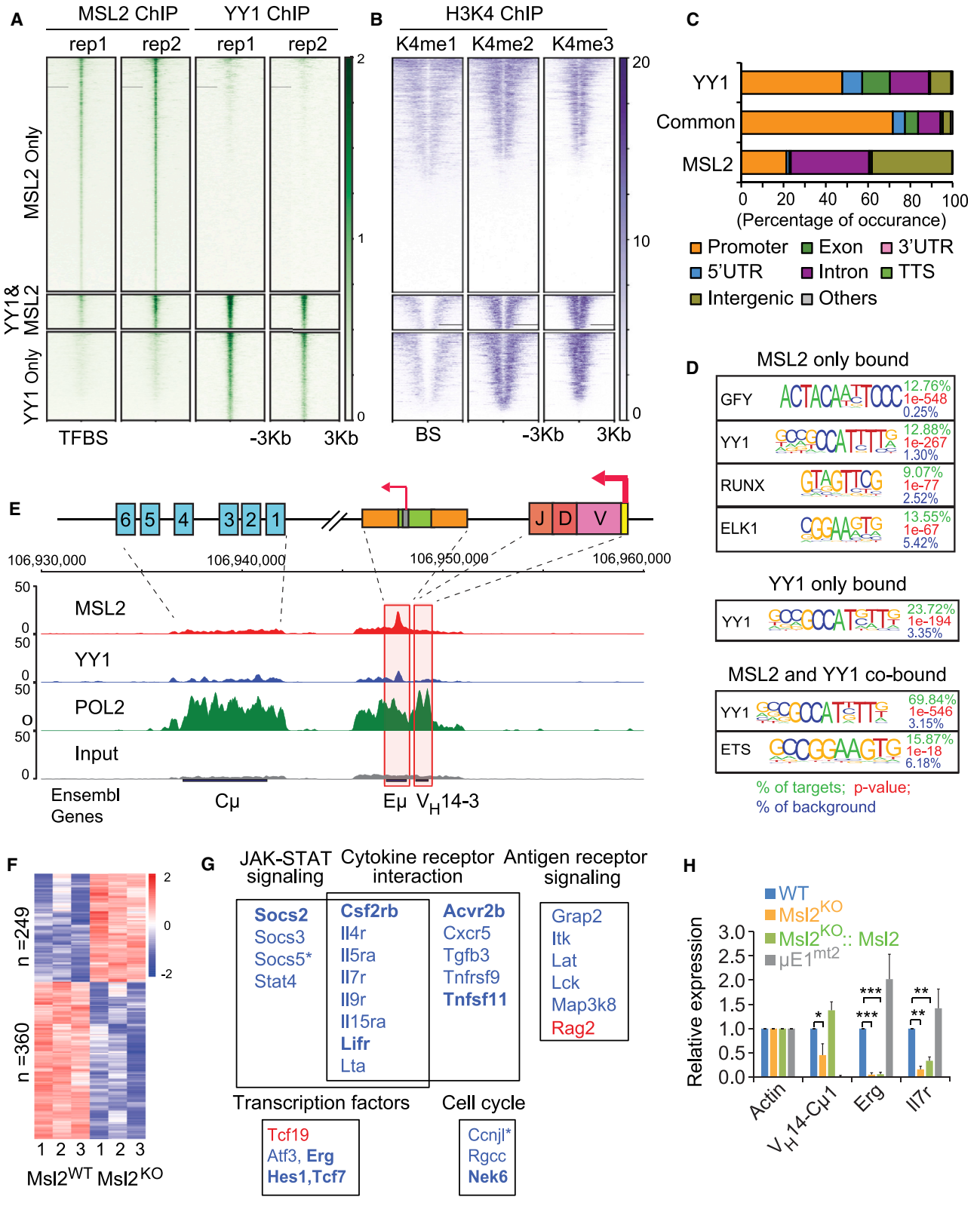
to *Msl2*^{WT} cells (Figure 4C). Moreover, we observed a reduction of interactions between the rearranged V_H14 promoter and Eμ as well as between V_H14 and 3′ RR in *Msl2*^{KO} cells relative to *Msl2*^{WT} cells (Figure 4C). As a control, no effects of the *Msl2* knockout were observed in the chromatin interactions between two sites in the *Rag* locus (Figure 4C). Thus, MSL2 may enhance chromatin loop formation at the μ locus via YY1-mediated recruitment at the μE1 box of the enhancer.

One of the prominent functions of the MSL/MOF complex is the acetylation of H4K16.[50] We observed an enrichment of H4K16ac at the transgenic V_H17.2.25 promoter in *Msl2*^{WT} pro-B cells that was diminished in *Msl2*^{KO} cells and at the promoter of the μE1^{mt1} transgene (Figures S4C and S4D). Likewise, we detected a statistically significant reduction of H4K16ac at the V_H14 promoter and Eμ core of the endogenous rearranged μ gene in *Slp65*^{KO}*Msl2*^{KO} pre-B cells relative to *Slp65*^{KO}*Msl2*^{WT} cells (Figure 4D). We also detected moderately but significantly reduced H4K16 acetylation after dTAG-induced YY1 degradation, whereas no significant changes were observed in H3K4me1 modification (Figure 4E). Moreover, no significant H4K16ac changes were detected at the β-actin locus.

To confirm the binding of the catalytic MOF subunit of the MSL/MOF complex at the Eμ enhancer, we performed a quantitative MOF ChIP analysis in *Slp65*^{KO} pre-B cells containing a rearranged μWT or μE1^{mt2} allele in *Slp65*^{KO}*Msl2*^{KO} pre-B cells carrying a μWT allele. This analysis indicated that MOF binding at the Eμ enhancer is reduced in cells lacking the μE1 box or an intact *Msl2* gene (Figure 5A). No change of MOF binding was detected at the *Phf8* gene body serving as a negative control. Taken together, these data suggest that MSL2 and the μE1 box of Eμ are required for recruiting enzymatic MOF activity, enhanced H4K16ac modification, and enhancer-driven chromatin looping.

MOF heterozygosity impairs B cell differentiation and μ gene expression

We further examined the role of the MSL/MOF complex *in vivo*. Since *Mof* homozygous knockout mice are embryonic lethal,[41,51] we analyzed mice carrying a heterozygous knockout of *Mof*, which is sufficient to decrease the overall levels of H4K16ac.[41] Flow cytometry analysis of the bone marrow of *Mof* $^{+/+}$ and *Mof* $^{+/-}$ mice indicated that numbers of pro-B, pre-B, and pre-pro-B cells were reduced in *Mof* $^{+/-}$ mice, whereas no significant changes were detected in common lymphoid progenitors (CLPs) (Figure 5B). To examine the effect of *Mof* heterozygosity on the expression of the endogenous μ gene, we performed RT-qPCR analysis on sorted pro-B cells and detected an approximately 2-fold decrease in the abundance of spliced Cμ transcripts (Figure 5C). Moreover, we detected a decrease of H4K16ac at the Eμ enhancer in *Mof* $^{+/-}$ pro-B cells (Figure 5D). We also analyzed the expression of distal J558 and proximal Q52 and V_H7183 segments in sorted pro-B cells by semi-quantitative RT-PCR. This analysis revealed a modest but uniform decrease in μ transcripts containing distal or proximal V_H sequences in *Mof*$^{+/-}$ cells relative to *Mof*$^{+/+}$ cells (Figure 5E). This result suggests that *Mof* heterozygosity affects μ gene expression but not the repertoire of rearranged V_H segments. Finally, we examined the expression of IgM on IgM^{lo}IgD^{hi} mature antigen-naive B cells and detected an ∼2-fold decrease in IgM expression in *Mof*$^{+/-}$ cells relative to *Mof*$^{+/+}$ cells (Figure 5F).



(legend on next page)

These data suggest that the catalytic activity of the MSL/MOF complex is required for optimal expression of the μ gene but not for the regulation of V_H gene rearrangements.

Genome-wide analysis of co-occupancy of YY1 and MSL2 in *Slp65*^{KO} pre-B cells

To examine the genome-wide occupancy of MSL2 and YY1 in *Slp65*^{KO} pre-B cells, we performed ChIP sequencing (ChIP-seq) analyses. By comparing MSL2 and YY1 occupancy, we found that 701 of 2,451 YY1-occupied sites overlapped with sites of MSL2 occupancy (Figure 6A; Table S6). To assess the overlap of MSL2-, YY1-, and MSL2- and YY1-bound sites with histone marks, we interrogated publicly available datasets of H3K4me1, me2, and me3 marks in *Rag1*$^{-/-}$ pro-B cells.[52] This analysis indicated that approximately 50% of MSL2-bound sites coincided with the H3K4me1 enhancer mark (Figure 6B). Bioinformatics analysis of the association of MSL2- and YY1-bound sites with gene loci revealed that the majority of YY1-MSL2 co-occupied sites are located in promoter regions (Figure 6C). Moreover, we examined the enrichment of TF motifs in regions of MSL2 and YY1 occupancy and MSL2-YY1 co-occupancy. This analysis indicated that the enrichment of the YY1 motif is even more pronounced at genomic sites of YY1-MSL2 co-occupancy than at sites of YY1 occupancy (Figure 6D). Sites of YY1-MSL2 co-occupancy also showed an enrichment of ETS family motifs. Additionally, we analyzed the genome-wide ChIP-seq data for the occupancy of the *Igh* locus by YY1, MSL2, and RNA Pol II and detected co-occupancy of MSL2 and YY1 at the Eμ enhancer (Figure 6E). The RNA Pol II-bound rearranged V_H14-3 promoter was weakly occupied by MSL2 and YY1 (Figure 6E), consistent with the weak association of YY1 with the V_H14 promoter in the ChIP analysis in the absence of a bifunctional crosslinker. At upstream unrearranged V_H segments, no MSL2 occupancy was detected (Figure 6E).

In addition to YY1-mediated enhancer-promoter looping, cohesin-mediated and CTCF-anchored chromosome looping allow for dynamic changes in chromosome folding.[53,54] We found that both MSL2 and YY1 are bound at the promoters of genes encoding CTCF; WAPL, a regulator of cohesin activity; RAD21, a structural component of the cohesin complex; and NIPBL, a member of the cohesin loading complex (Figures S5A–S5D). The binding of YY1 and MSL2 at the *Ctcf* and *Rad21* promoters was further corroborated by quantitative ChIP analysis (Figures S5E and S5F). Subsequently, we analyzed the effects of *Msl2* knockout on the expression of *Ctcf* and found a modest but significant decrease in *Msl2*^{KO} pre-B cells relative to *Msl2*^{WT} pre-B cells (Figure S5G). This decrease was rescued by the ectopic expression of MSL2 in *Msl2*^{KO} cells. Taken together, our data support a model where YY1-mediated recruitment of the MSL/MOF complex at specific genes may augment chromatin looping and enhancer-promoter interactions.

MSL2 deficiency results in altered expression of regulators of early B cell differentiation

We also conducted genome-wide RNA sequencing (RNA-seq) analysis to identify genes that are deregulated in *Slp65*^{KO}*Msl2*^{KO} pre-B cells versus *Slp65*^{KO}*Msl2*^{WT} pre-B cells. In this analysis, we identified 249 genes that were upregulated in the *Msl2* mutant cells and 360 downregulated genes (Figures 6F; Table S7). In the set of downregulated genes, we identified many genes in the JAK-STAT, cytokine receptor, and antigen receptor signaling pathways (Figure 6G). In addition, we found several genes encoding TFs and some cell cycle regulators. Among the genes that were downregulated in *Slp65*^{KO}*Msl2*^{KO} pre-B cells, we identified *Hexim2*, which exhibited co-occupancy by MSL2 and YY1 in its promoter region (Figure S5H; Table S8). Consequently, the loss of MSL2 and dTAG13-induced degradation of YY1 resulted in reduced expression of *Hexim2* (Figures S5I and S5J). By interrogating the RNA-seq and ChIP-seq datasets, we also uncovered deregulated genes that are occupied by MSL2 in WT cells, suggesting a direct role of MSL2 in their regulation. In particular, we identified *Erg*, encoding an ETS-related

Figure 6. Genome-wide analysis of MSL2 and YY1 binding and mRNA expression

(A) Heatmap of MSL2 and YY1 ChIP signal 3 kb from peak center in 2 replicates of *Slp65*^{KO} pre-B cells. The peaks are clustered as MSL2 bound only, MSL2 and YY1 co-bound, and YY1 bound only. The count per million (CPM) normalized value of the ChIP signal is scaled to the intensity of the heatmap signal. The data represent two biological replicates.

(B) Heatmap of histone H3K4 monomethyl (me1), dimethyl (me2), and trimethyl (me3) ChIP signals ± 3 kb from peak center (*Rag1*$^{-/-}$ pro-B cells).[52] The peaks are clustered as MSL2 bound only, MSL2 and YY1 co-bound, and YY1 bound only. The CPM normalized value of the ChIP signal is scaled to the intensity of the heatmap signal.

(C) Distribution of MSL2-only, common, and YY1-only bound regions in different compartments of the genome (promoter, exon, 3′ UTR [untranslated region], 5′ UTR, intron, intergenic, and TTS [transcription termination site]).

(D) Significantly enriched *de novo* motifs in repeat-masked MSL2-only, YY1-only, and co-bound peaks in *Slp65*^{KO} pre-B cells. The *de novo* motifs are assigned to the relevant TF binding site predicted by the HOMER algorithm. The percentage of occurrence (green), *p* value (red), and percentage of occurrence in the background (blue) for each motif are depicted.

(E) Distribution of YY1, MSL2, and RNA POL II binding signals at the *Igh* locus (OW971649.1). The graphical representation of the *Igh* locus (top) and the ChIP signals on the gene body (Cμ, 106,937,061–106,941,237), the Eμ enhancer (106,947,187–106,948,179), and the V_H14 region (106,948,788–106,949,265) are depicted. The y axis indicates the CPM normalized value of the ChIP signal for MSL2, YY1, RNA POL II, and input.

(F) Heatmap of differentially expressed genes (1.5-fold cutoff) between *Msl2*^{KO} and *Msl2*^{WT} pre-B cells. The number of up- or downregulated genes is indicated (left). The log CPM normalized read counts are scaled to the intensity of the heatmap signal (*Z* score). Each condition includes 3 replicates.

(G) List of genes encoding components of JAK-STAT signaling, cytokine receptor interaction, antigen receptor signaling, TFs, and cell cycle pathways that are upregulate (red) or downregulated (blue) in *Msl2*^{KO} cells. *, genes bound by MSL2.

(H) RT-qPCR analysis to assess mRNA expression of the VDJ-rearranged μ gene using V_H14-Cμ exon 1-spanning primers and transcripts of the *Erg* and *Il7ra* genes in *Msl2*^{WT}, *Msl2*^{KO} and *Msl2*^{KO}::*Msl2 Slp65*^{KO} pre-B cells. For comparison, expression of these genes was also determined in *Slp65*^{KO} pre-B cells carrying the μE1^{mt2} mutation in the VDJ-rearranged μ gene. Error bars represent the standard deviation of three biological replicates. Statistical significance between WT and mutant cells was measured by an unpaired one-tailed Student's t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TF that regulates VDJ recombination and early B cell differentiation,[55] as an MSL2-bound gene (Figure S6A). *Erg* is markedly downregulated in *Msl2*^{KO} versus *Msl2*^{WT} pre-B cells, but, in contrast to *Igh*, it is not re-expressed upon forced MSL2 expression (Figures 6H and S6A). No significant effect of the μE1^{mt2} mutation on *Erg* expression was observed (Figure 6H). The transcription factor ERG has been implicated in the regulation of *Il7ra* expression,[56] and, consistent with this notion, the number of *Il7ra* transcripts was reduced in *Msl2*^{KO} versus *Msl2*^{WT} pre-B cells (Figures 6G and 6H). Thus, the lack of an effect of forced MSL2 expression on *Erg* may be accounted for by a feedback loop between *Erg* and *Il7ra* that is disrupted by the accelerated differentiation (see below) and reduced IL-7 signaling.

Early B cell differentiation is regulated by an interplay between pre-BCR and IL-7R signaling, in which reduced IL-7R signaling results in upregulation of *Rag* gene expression, increased Ig kappa light chain gene expression and rearrangement, as well as reduced cell proliferation.[8,57] In *Msl2*^{KO} pre-B cells, we observed diminished phosphorylation of the nuclear effector of IL-7R signaling, STAT5 (Figure S6B), and an increase in the expression of *Rag2* and the constant region of the *Igk* and *Igl* genes (Figures 6G and S6C). Concomitantly, we found that cell proliferation of *Msl2*^{KO} pre-B cells was reduced relative to *Msl2*^{WT} cells (Figure S6D). Notably, the reduced proliferation and attenuated IL-7 receptor signaling of *Msl2*^{KO} pre-B cells is consistent with the observed decrease in the expression of CDK6 and Cyclin D3 (Figure S6E). Taken together, these data suggest that MSL2/MOF constitutes an important regulator of early B cell differentiation by affecting the expression of many genes involved in the IL-7R signaling pathway, Ig gene expression and recombination, as well as cell proliferation.

DISCUSSION

In this study, we provide a comprehensive and innovative characterization of the lymphocyte-specific immunoglobulin Eμ enhancer proteome, enabling the detection of both direct TF binding and indirect association of enhancer-binding proteins, such as chromatin modifiers. By employing a reverse DNA sequence polarity approach, which globally interferes with sequence-specific TF binding, we determined the entire enhancer proteome. Moreover, the analysis of the proteome of the Eμ enhancer carrying mutations in the μE1 box allowed for the identification of components of the MSL/MOF complex that are predominantly recruited to the core enhancer via interactions with the YY1 TF. By functional assays with transgenic and endogenous *Igh* genes carrying mutations in the YY1 binding site and by knocking out the *Msl2* gene or degrading the YY1 protein in primary *Slp65*^{KO} pre-B cells, we found that MSL2 regulates *Igh* expression in pro-B and pre-B cells. In addition, we identified a set of YY1-bound genomic sites that are co-occupied by MSL2 and are associated with genes that are deregulated in *Msl2* knockout pre-B cells.

The MSL/MOF complex is a well-studied epigenetic modulator of gene expression that regulates gene dosage compensation in *D. melanogaster*.[58] In this context, the MSL/MOF complex binds to the male X chromosome via GA repeats or extended Pion X sites.[38,59–62] Moreover, binding of MSL2 to GA-rich motifs can be enhanced by an interaction of MSL2 with CLAMP.[63] In contrast, the mammalian MSL/MOF complex is bound at many autosomal sites that are enriched for $(CAGA)_n$ motifs.[38,44] Our MSL2 ChIP-seq analysis in pre-B cells identified a large set of MSL2-occupied sites that are enriched in $(CAGA)_n$ motifs if no repeat masker is used for analysis (data not shown). In addition, a specific set of MSL2-bound sites overlapped with occupancy by YY1, suggesting that MSL2 is recruited to this set of genomic sites via interactions with YY1. This YY1-mediated targeting of MSL2 is also supported by our finding that MSL2 occupancy at the Eμ core enhancer is abrogated upon dTAG13-induced degradation of YY1.

YY1 has been shown to interact with multiple Ig enhancers, including the intronic Eμ and the 3′ RR enhancers of the *Igh* locus and the *Igk* 3′ enhancer.[18,22,23,64] Accordingly, YY1 represents a key regulator of early B cell differentiation and the germinal center developmental program.[45,65] Molecular analysis indicated that YY1 plays an important role in the formation of long-distance DNA loops between the Eμ and 3′ RR as well as in DNA loops between the Eμ region and V_H promoters.[18,23,66] Recently, YY1 has been shown to regulate enhancer-promoter interactions by dimer interactions between enhancer- and promoter-bound YY1 and by forming loops within CTCF-mediated chromatin topologically associated domains (TADs).[47] Gene activation and repression by YY1 involves interactions with INO80-containing chromatin remodeling complex and polycomb group proteins, respectively.[67–69]

In this study, we identified the MSL/MOF complex as an additional co-regulator of YY1 that augments DNA looping between the Eμ and 3′ RR enhancers and interactions between the enhancers and the rearranged V_H promoter. In particular, the knockout of the *Msl2* gene in pre-B cells impairs chromatin interactions and reduces the expression of an endogenous rearranged μ allele approximately 2-fold. Similarly, we observed a 2-fold decrease in the expression of unrearranged germline Cμ transcripts in *Rag2*-deficient *Msl2*^{KO} pro-B cells, consistent with the role of Eμ and the 3′ RR interaction in regulating the expression of sterile Ig germline transcripts.[70,71] Notably, we detected co-occupancy of YY1 and MSL2 at several genes encoding components of the CTCF/cohesion machinery of 3D chromosome architecture, including CTCF, RAD21, and WAPL, raising the interesting possibility that YY1 targeting of the MSL2/MOF complex may also affect CTCF/cohesion-mediated DNA looping. Hence, our data suggest that MSL2/MOF serves as a regulatory component in *Igh* locus organization.

Although YY1 has been described as an E1 box-binding protein, we noted a marked difference in the transcriptional effects of the YY1 protein degradation and the E1 box mutations, suggesting that additional factors interact with the E1 box. In the proteomics analysis of proteins differentially bound at the μWT enhancer vs. μE1^{mt1} enhancer, we identified multiple members of the ZHX family of TFs. These proteins contain multiple Zn fingers and homeodomains, and they bind 5′ CCANCAC and 5′ GGNCAACA motifs, which partially overlap with the E1 box.[72] Insight into the function of these proteins is limited, but ZHX2 has been shown to regulate natural killer cell maturation.[73] In addition, we identified ZFP383 as an E1 box-binding protein that shares around 50% sequence identity in the DNA-binding domain with YY1. In our functional analysis of the E1 box

mutation, we observed a marked defect in the expression of rearranged *Igh* alleles in transgenic pro-B cells and *Slp65*^{KO} pre-B cells but only a modest defect in the expression of unrearranged alleles in *Rag2*-deficient pro-B cells. Thus, E1 box-dependent but YY1:MSL2/MOF-independent regulation of the Eμ enhancer may occur at subsequent developmental stages and/or involve other TFs or non-coding RNA. In our proteomics screen, we additionally identified an E1 box-dependent enrichment of multiple components of the splicing machinery, which has been shown to markedly enhance transcription elongation.[74,75] Sequential recruitment of DNA-binding proteins and associated protein complexes has been found to modulate gene expression during cell differentiation and development.[76]

In addition to the regulation of the *Igh* locus by MSL2, we identified several genes that are downregulated in *Msl2*^{KO} pre-B cells. However, relatively few down-regulated genes were also bound by MSL2, identifying them as potential direct targets. In *Msl2*^{KO} pre-B cells, we observed a marked downregulation of the MSL2-bound *Erg* gene, encoding a TF that regulates early B cell differentiation, *Il7ra* expression, and *Igh* gene rearrangement.[55,56] ERG binds the μA box of the Eμ enhancer, and ERG1-occupied sites at other genes show a strong enrichment of motifs for EBF1, PAX3, E2A, and FOXO1, suggesting that ERG cooperates with other B cell TFs in regulating B cell-specific gene expression.[55] In *Msl2*^{KO} pre-B cells, we observed downregulation of *Il7ra* expression, which may account for the diminished cell proliferation and upregulation of *Rag* expression. In addition, we detected increased numbers of C region transcripts of the Ig kappa and lambda loci, suggesting that the *Msl2* deficiency in pre-B cells results in enhanced differentiation and reduced proliferation. In contrast, we did not observe these phenotypes in μE1^{mt2} pre-B cells, in which μ transcription and μ protein expression are also decreased. Thus, the MSL/MOF complex may specifically regulate early B cell differentiation by affecting the expression of genes involved in pre-BCR and IL-7R signaling.

Limitations of the study

Various mechanisms could account for the function of the YY1-associated MSL/MOF complex in affecting V_H promoter activity and μ gene transcription. MSL/MOF could directly affect the interaction of enhancer-bound proteins with promoter-bound proteins, changing the chromatin landscape via MOF-mediated histone acetylation or indirectly via changing the activity of YY1. In addition to H4K16 acetylation, MOF has been shown to acetylate other nuclear proteins, including Lamin A and the TFs p53 and NRF2.[77–79] Future experiments will have to determine whether the role of YY1-associated MSL/MOF in regulating enhancer-promoter interaction involves an acetylation of YY1 or other enhancer-bound proteins.

The MSL/MOF complex is best characterized as an important regulator of gene dosage compensation.[58] Recent analysis of MSL2-deficient neuronal progenitor cells carrying distinguishable alleles indicated a role of MSL2 in preserving biallelic expression of specific dosage-sensitive genes sets.[80] Antigen receptor genes are well studied paradigms of mitotically stable random monoallelic gene expression. In particular, the analysis of B lineage cells carrying distinguishable *Igk* alleles has shown that allelic differences in asynchronous replication timing patterns are established in early progenitors and stably maintained in mature B cells.[81] Moreover, in case of the immunoglobulin heavy and light gene loci, the stability of monoallelic expression is further enforced by the process of allelic exclusion.[6] Therefore, an analysis of B-lineage cells with distinguishable alleles will be necessary to determine whether the MSL2/MOF complex also plays a role in allele-specific immunoglobulin gene expression.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Mice
  - Cell lines
  - Primary cell cultures
- METHOD DETAILS
  - Generation of transgenic cell lines
  - CRISPR/Cas9 editing of *Slp65*^{KO} pre-B cells
  - Generation of μE1^{mt3} pro-B cell lines
  - Cell cultures and SILAC labeling
  - Intracellular FACS
  - Preparation of nuclei
  - qRT-PCR analysis of mRNA expression
  - Co-immunoprecipitation (CoIP)
  - Chromatin immunoprecipitation (ChIP)
  - Native chromatin immunoprecipitation (native ChIP)
  - DNA affinity chromatography experiments
  - Proteomics
  - Formaldehyde-assisted isolation of regulatory elements (FAIRE)
  - 3C assay
  - RNA-seq analysis
  - ChIP-seq analysis
  - Analysis of publicly available H3K4 methylation ChIP-seq data
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114456.

ACKNOWLEDGMENTS

We thank I. Falk, G. Nerz, and H.-J. Schwarz for technical assistance and the MPI core facilities for Proteomics (Ida Suppanz and Daniel Eilertz), Deep Sequencing, Mouse Transgenesis, and Flow Cytometry for invaluable support. We are grateful to Kyunjin Boo for help with 3C experiments, Eirini Trompouki for critical reading of the manuscript, and members of the Akhtar, Sen, and Grosschedl labs for discussions. This work was supported by funds of the Max Planck Society (Grosschedl, Mittler, and Akhtar labs) and by intramural research funds of the NIA (Sen lab). Y.P. was partially supported by the Thai government scholarship from the Development and Promotion of Science and Technology Talents Project and by the Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic.

AUTHOR CONTRIBUTIONS

Y.P., R.G., and G.M. designed experiments. Y.P. performed DNA pull-down, coIP, mRNA expression, ChIP, immunoblotting, 3C, and experiments and generated CRISPR-Cas9-modified cell lines. F.-Z.B. and F.M. conducted

experiments relating μE1^{mt3} mice. C.P.-R. performed MOF ChIP and analyzed *Mof* $^{+/-}$ mice by FACS and RT-qPCR. H.R. performed coIP and MSL2 ChIP-seq experiments. S.R. and P.C. carried out bioinformatics analysis. G.M. performed and analyzed proteomics experiments. A.G. analyzed FACS experiments. Conceptualization, Y.P., G.M., and R.G.; writing, Y.P. and R.G. with input from G.M.; supervision, A.A. (analysis of *Mof*$^{+/-}$ mice and MOF and MSL2 data), R.S. (analysis of μE1^{mt3} mice and 3C data), G.M. (proteomics data analysis), and R.G. (overall supervision).

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 29, 2023
Revised: May 2, 2024
Accepted: June 21, 2024
Published: July 9, 2024

REFERENCES

1. Jung, D., and Alt, F.W. (2004). Unraveling V(D)J Recombination: Insights into Gene Regulation. Cell *116*, 299–311.
2. Schatz, D.G., and Ji, Y. (2011). Recombination centres and the orchestration of V(D)J recombination. Nat. Rev. Immunol. *11*, 251–263.
3. Vettermann, C., and Schlissel, M.S. (2010). Allelic exclusion of immunoglobulin genes: models and mechanisms. Immunol. Rev. *237*, 22–42.
4. Bergman, Y., and Cedar, H. (2004). A stepwise epigenetic process controls immunoglobulin allelic exclusion. Nat. Rev. Immunol. *4*, 753–761.
5. Krangel, M.S. (2009). Mechanics of T cell receptor gene rearrangement. Curr. Opin. Immunol. *21*, 133–139.
6. Levin-Klein, R., and Bergman, Y. (2014). Epigenetic regulation of monoallelic rearrangement (allelic exclusion) of antigen receptor genes. Front. Immunol. *5*, 625.
7. Baizan-Edge, A., Stubbs, B.A., Stubbington, M.J.T., Bolland, D.J., Tabbada, K., Andrews, S., and Corcoran, A.E. (2021). IL-7R signaling activates widespread V(H) and D(H) gene usage to drive antibody diversity in bone marrow B cells. Cell Rep. *36*, 109349.
8. Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. Nat. Rev. Immunol. *14*, 69–80.
9. Corcoran, A.E., Riddell, A., Krooshoop, D., and Venkitaraman, A.R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature *391*, 904–907.
10. Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell *33*, 729–740.
11. Gillies, S.D., Morrison, S.L., Oi, V.T., and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell *33*, 717–728.
12. Jenuwein, T., and Grosschedl, R. (1991). Complex pattern of immunoglobulin mu gene expression in normal and transgenic mice: nonoverlapping regulatory sequences govern distinct tissue specificities. Genes Dev. *5*, 932–943.
13. Jenuwein, T., Forrester, W.C., Fernández-Herrero, L.A., Laible, G., Dull, M., and Grosschedl, R. (1997). Extension of chromatin accessibility by nuclear matrix attachment regions. Nature *385*, 269–272.
14. Chen, J., Young, F., Bottaro, A., Stewart, V., Smith, R.K., and Alt, F.W. (1993). Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. EMBO J. *12*, 4635–4645.
15. Sakai, E., Bottaro, A., Davidson, L., Sleckman, B.P., and Alt, F.W. (1999). Recombination and transcription of the endogenous Ig heavy chain locus is effected by the Ig heavy chain intronic enhancer core region in the absence of the matrix attachment regions. Proc. Natl. Acad. Sci. USA *96*, 1526–1531.
16. Perlot, T., Alt, F.W., Bassing, C.H., Suh, H., and Pinaud, E. (2005). Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc. Natl. Acad. Sci. USA *102*, 14362–14367.
17. Chakraborty, T., Perlot, T., Subrahmanyam, R., Jani, A., Goff, P.H., Zhang, Y., Ivanova, I., Alt, F.W., and Sen, R. (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.
18. Guo, C., Gerasimova, T., Hao, H., Ivanova, I., Chakraborty, T., Selimyan, R., Oltz, E.M., and Sen, R. (2011). Two Forms of Loops Generate the Chromatin Conformation of the Immunoglobulin Heavy-Chain Gene Locus. Cell *147*, 332–343.
19. Marquet, M., Garot, A., Bender, S., Carrion, C., Rouaud, P., Lecardeur, S., Denizot, Y., Cogné, M., and Pinaud, E. (2014). The Eμ Enhancer Region Influences H Chain Expression and B Cell Fate without Impacting IgVH Repertoire and Immune Response In Vivo. J. Immunol. *193*, 1171–1183.
20. Zhang, X., Zhang, Y., Ba, Z., Kyritsis, N., Casellas, R., and Alt, F.W. (2019). Fundamental roles of chromatin loop extrusion in antibody class switching. Nature *575*, 385–389.
21. Ernst, P., and Smale, S.T. (1995). Combinatorial regulation of transcription II: the immunoglobulin μ heavy chain gene. Immunity *2*, 427–438.
22. Park, K., and Atchison, M.L. (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. Medvedovic, J., Ebert, A., Tagoh, H., Tamir, I.M., Schwickert, T.A., Novatchkova, M., Sun, Q., Huis In 't Veld, P.J., Guo, C., Yoon, H.S., et al. (2013). Flexible Long-Range Loops in the VH Gene Region of the Igh Locus Facilitate the Generation of a Diverse Antibody Repertoire. Immunity *39*, 229–244.
24. Jenuwein, T., Forrester, W.C., Qiu, R.G., and Grosschedl, R. (1993). The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. Genes Dev. *7*, 2016–2032.
25. Forrester, W.C., Fernández, L.A., and Grosschedl, R. (1999). Nuclear matrix attachment regions antagonize methylation-dependent repression of long-range enhancer-promoter interactions. Genes Dev. *13*, 3003–3014.
26. Fernández, L.A., Winkler, M., and Grosschedl, R. (2001). Matrix attachment region-dependent function of the immunoglobulin mu enhancer involves histone acetylation at a distance without changes in enhancer occupancy. Mol. Cell Biol. *21*, 196–208.
27. Mittler, G., Butter, F., and Mann, M. (2009). A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. Genome Res. *19*, 284–293.
28. Rusk, N. (2009). Reverse ChIP. Nat. Methods *6*, 187.
29. Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Côté, J., and Lucchesi, J.C. (2005). A Human Protein Complex Homologous to the Drosophila MSL Complex Is Responsible for the Majority of Histone H4 Acetylation at Lysine 16. Mol. Cell Biol. *25*, 9175–9188.
30. Conrad, T., and Akhtar, A. (2012). Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nat. Rev. Genet. *13*, 123–134.
31. Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., et al. (2006). Nuclear Pore Components Are Involved in the Transcriptional Regulation of Dosage Compensation in Drosophila. Mol. Cell *21*, 811–823.
32. Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide Analysis Reveals MOF as a Key Regulator of Dosage Compensation and Gene Expression in Drosophila. Cell *133*, 813–828.
33. Gaub, A., Sheikh, B.N., Basilicata, M.F., Vincent, M., Nizon, M., Colson, C., Bird, M.J., Bradner, J.E., Thevenon, J., Boutros, M., and Akhtar, A. (2020). Evolutionary conserved NSL complex/BRD4 axis controls transcription activation via histone acetylation. Nat. Commun. *11*, 2243.

34. Lam, K.C., Mühlpfordt, F., Vaquerizas, J.M., Raja, S.J., Holz, H., Luscombe, N.M., Manke, T., and Akhtar, A. (2012). The NSL Complex Regulates Housekeeping Genes in Drosophila. PLoS Genet. *8*, 10027366–e1002818.

35. Prestel, M., Feller, C., and Becker, P.B. (2010). Dosage compensation and the global re-balancing of aneuploid genomes. Genome Biol. *11*, 216.

36. Sheikh, B.N., Guhathakurta, S., and Akhtar, A. (2019). The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis. EMBO Rep. *20*, e47630.

37. Valsecchi, C.I.K., Basilicata, M.F., Semplicio, G., Georgiev, P., Gutierrez, N.M., and Akhtar, A. (2018). Facultative dosage compensation of developmental genes on autosomes in Drosophila and mouse embryonic stem cells. Nat. Commun. *9*, 3626.

38. Chelmicki, T., Dündar, F., Turley, M.J., Khanam, T., Aktas, T., Ramírez, F., Gendrel, A.V., Wright, P.R., Videm, P., Backofen, R., et al. (2014). MOF-associated complexes ensure stem cell identity and *Xist* repression. Elife *3*, e02024.

39. Gupta, A., Hunt, C.R., Pandita, R.K., Pae, J., Komal, K., Singh, M., Shay, J.W., Kumar, R., Ariizumi, K., Horikoshi, N., et al. (2013). T-cell-specific deletion of Mof blocks their differentiation and results in genomic instability in mice. Mutagenesis *28*, 263–270.

40. Valerio, D.G., Xu, H., Eisold, M.E., Woolthuis, C.M., Pandita, T.K., and Armstrong, S.A. (2017). Histone acetyltransferase activity of MOF is required for adult but not early fetal hematopoiesis in mice. Blood *129*, 48–59.

41. Pessoa Rodrigues, C., Herman, J.S., Herquel, B., Valsecchi, C.I.K., Stehle, T., Grün, D., and Akhtar, A. (2020). Temporal expression of MOF acetyltransferase primes transcription factor networks for erythroid fate. Sci. Adv. *6*, eaaz4815.

42. Jumaa, H., Mitterer, M., Reth, M., and Nielsen, P.J. (2001). The absence of SLP65 and Btk blocks B cell development at the preB cell receptor-positive stage. Eur. J. Immunol. *31*, 2164–2169.

43. Ta, V.B.T., de Bruijn, M.J.W., Matheson, L., Zoller, M., Bach, M.P., Wardemann, H., Jumaa, H., Corcoran, A., and Hendriks, R.W. (2012). Highly Restricted Usage of Ig H Chain VH14 Family Gene Segments in Slp65-Deficient Pre-B Cell Leukemia in Mice. J. Immunol. *189*, 4842–4851.

44. Valsecchi, C.I.K., Basilicata, M.F., Georgiev, P., Gaub, A., Seyfferth, J., Kulkarni, T., Panhale, A., Semplicio, G., Manjunath, V., Holz, H., et al. (2021). RNA nucleation by MSL2 induces selective X chromosome compartmentalization. Nature *589*, 137–142.

45. Liu, H., Schmidt-Supprian, M., Shi, Y., Hobeika, E., Barteneva, N., Jumaa, H., Pelanda, R., Reth, M., Skok, J., Rajewsky, K., and Shi, Y. (2007). Yin Yang 1 is a critical regulator of B-cell development. Genes Dev. *21*, 1179–1189.

46. Winter, G.E., Buckley, D.L., Paulk, J., Roberts, J.M., Souza, A., Dhe-Paganon, S., and Bradner, J.E. (2015). DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. Science *348*, 1376–1381.

47. Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573–1588.e28.

48. Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W., and Frock, R.L. (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification–mediated high-throughput genome-wide translocation sequencing. Nat. Protoc. *11*, 853–871.

49. Hao, B., Naik, A.K., Watanabe, A., Tanaka, H., Chen, L., Richards, H.W., Kondo, M., Taniuchi, I., Kohwi, Y., Kohwi-Shigematsu, T., and Krangel, M.S. (2015). An anti-silencer– and SATB1-dependent chromatin hub regulates Rag1 and Rag2 gene expression during thymocyte development. J. Exp. Med. *212*, 809–824.

50. Conrad, T., Cavalli, F.M.G., Holz, H., Hallacli, E., Kind, J., Ilik, I., Vaquerizas, J.M., Luscombe, N.M., and Akhtar, A. (2012). The MOF Chromobarrel Domain Controls Genome-wide H4K16 Acetylation and Spreading of the MSL Complex. Dev. Cell *22*, 610–624.

51. Thomas, T., Dixon, M.P., Kueh, A.J., and Voss, A.K. (2008). Mof (MYST1 or KAT8) Is Essential for Progression of Embryonic Development Past the Blastocyst Stage and Required for Normal Chromatin Architecture. Mol. Cell Biol. *28*, 5093–5105.

52. Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J., et al. (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. Nat. Immunol. *11*, 635–643.

53. Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., Mach, P., Kryzhanovska, M., Tihanyi, G., Kohler, H., et al. (2022). Nonlinear control of transcription through enhancer–promoter interactions. Nature *604*, 571–577.

54. Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer–promoter contacts in gene expression control. Nat. Rev. Genet. *20*, 437–455.

55. Ng, A.P., Coughlan, H.D., Hediyeh-Zadeh, S., Behrens, K., Johanson, T.M., Low, M.S.Y., Bell, C.C., Gilan, O., Chan, Y.C., Kueh, A.J., et al. (2020). An Erg-driven transcriptional program controls B cell lymphopoiesis. Nat. Commun. *11*, 3013.

56. Søndergaard, E., Rauch, A., Michaut, M., Rapin, N., Rehn, M., Wilhelmson, A.S., Camponeschi, A., Hasemann, M.S., Bagger, F.O., Jendholm, J., et al. (2019). ERG Controls B Cell Development by Promoting Igh V-to-DJ Recombination. Cell Rep. *29*, 2756–2769.e6.

57. Herzog, S., Reth, M., and Jumaa, H. (2009). Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. Nat. Rev. Immunol. *9*, 195–205.

58. Samata, M., and Akhtar, A. (2018). Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. Annu. Rev. Biochem. *87*, 323–350.

59. Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J., and Kuroda, M.I. (2008). A Sequence Motif within Chromatin Entry Sites Directs MSL Establishment on the Drosophila X Chromosome. Cell *134*, 599–609.

60. Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. (2006). Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes Dev. *20*, 858–870.

61. Straub, T., Zabel, A., Gilfillan, G.D., Feller, C., and Becker, P.B. (2013). Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. Genome Res. *23*, 473–485.

62. Villa, R., Schauer, T., Smialowski, P., Straub, T., and Becker, P.B. (2016). PionX sites mark the X chromosome for dosage compensation. Nature *537*, 244–248.

63. Albig, C., Tikhonova, E., Krause, S., Maksimenko, O., Regnard, C., and Becker, P.B. (2019). Factor cooperation for chromosome discrimination in Drosophila. Nucleic Acids Res. *47*, 1706–1724.

64. Gordon, S.J., Saleque, S., and Birshtein, B.K. (2003). Yin Yang 1 is a lipopolysaccharide-inducible activator of the murine 3' Igh enhancer, hs3. J. Immunol. *170*, 5549–5557.

65. Green, M.R., Monti, S., Dalla-Favera, R., Pasqualucci, L., Walsh, N.C., Schmidt-Supprian, M., Kutok, J.L., Rodig, S.J., Neuberg, D.S., Rajewsky, K., et al. (2011). Signatures of murine B-cell development implicate Yy1 as a regulator of the germinal center-specific program. Proc. Natl. Acad. Sci. USA *108*, 2873–2878.

66. Mehra, P., Gerasimova, T., Basu, A., Jha, V., Banerjee, A., Sindhava, V., Gray, F., Berry, C.T., Sen, R., and Atchison, M.L. (2016). YY1 controls Eμ-3'RR DNA loop formation and immunoglobulin heavy chain class switch recombination. Blood Adv. *1*, 15–20.

67. Wilkinson, F.H., Park, K., and Atchison, M.L. (2006). Polycomb recruitment to DNA *in vivo* by the YY1 REPO domain. Proc. Natl. Acad. Sci. USA *103*, 19296–19301.

68. Wu, S., Murai, S., Kataoka, K., and Miyagishi, M. (2008). Yin Yang 1 induces transcriptional activity of p73 through cooperation with E2F1. Biochem. Biophys. Res. Commun. *365*, 75–81.

69. Cai, Y., Jin, J., Yao, T., Gottschalk, A.J., Swanson, S.K., Wu, S., Shi, Y., Washburn, M.P., Florens, L., Conaway, R.C., and Conaway, J.W. (2007). YY1 functions with INO80 to activate transcription. Nat. Struct. Mol. Biol. *14*, 872–874.

70. Li, F., and Eckhardt, L.A. (2009). A role for the IgH intronic enhancer Eμ in enforcing allelic exclusion. J. Exp. Med. *206*, 153–167.

71. Braikia, F.-Z., Conte, C., Moutahir, M., Denizot, Y., Cogné, M., and Khamlichi, A.A. (2015). Developmental Switch in the Transcriptional Activity of a Long-Range Regulatory Element. Mol. Cell Biol. *35*, 3370–3380.

72. Zhu, L., Ding, R., Yan, H., Zhang, J., and Lin, Z. (2020). ZHX2 drives cell growth and migration via activating MEK/ERK signal and induces Sunitinib resistance by regulating the autophagy in clear cell Renal Cell Carcinoma. Cell Death Dis. *11*, 337.

73. Tan, S., Guo, X., Li, M., Wang, T., Wang, Z., Li, C., Wu, Z., Li, N., Gao, L., Liang, X., and Ma, C. (2021). Transcription factor Zhx2 restricts NK cell maturation and suppresses their antitumor immunity. J. Exp. Med. *218*, e20210009.

74. Fong, Y.W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. Nature *414*, 929–933.

75. Yu, Y., and Reed, R. (2015). FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. Proc. Natl. Acad. Sci. USA *112*, 8608–8613.

76. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell *38*, 576–589.

77. Karoutas, A., Szymanski, W., Rausch, T., Guhathakurta, S., Rog-Zielinska, E.A., Peyronnet, R., Seyfferth, J., Chen, H.R., de Leeuw, R., Herquel, B., et al. (2019). The NSL complex maintains nuclear architecture stability via lamin A/C acetylation. Nat. Cell Biol. *21*, 1248–1260.

78. Sykes, S.M., Stanek, T.J., Frank, A., Murphy, M.E., and McMahon, S.B. (2009). Acetylation of the DNA Binding Domain Regulates Transcription-independent Apoptosis by p53. J. Biol. Chem. *284*, 20197–20205.

79. Chen, Z., Ye, X., Tang, N., Shen, S., Li, Z., Niu, X., Lu, S., and Xu, L. (2014). The histone acetylranseferase hMOF acetylates Nrf2 and regulates anti-drug responses in human non-small cell lung cancer. Br. J. Pharmacol. *171*, 3196–3211.

80. Sun, Y., Wiese, M., Hmadi, R., Karayol, R., Seyfferth, J., Martinez Greene, J.A., Erdogdu, N.U., Deboutte, W., Arrigoni, L., Holz, H., et al. (2023). MSL2 ensures biallelic gene expression in mammals. Nature *624*, 173–181.

81. Farago, M., Rosenbluh, C., Tevlin, M., Fraenkel, S., Schlesinger, S., Masika, H., Gouzman, M., Teng, G., Schatz, D., Rais, Y., et al. (2012). Clonal allelic predetermination of immunoglobulin-κ rearrangement. Nature *490*, 561–565.

82. Alt, F.W., Yancopoulos, G.D., Blackwell, T.K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S., and Baltimore, D. (1984). Ordered rearrangement of immunoglobulin heavy chain variable region segments. EMBO J. *3*, 1209–1219. https://doi.org/10.1002/j.1460-2075.1984.tb01955.x.

83. Antonio Urrutia, G., Ramachandran, H., Cauchy, P., Boo, K., Ramamoorthy, S., Boller, S., Dogan, E., Clapes, T., Trompouki, E., Torres-Padilla, M.E., et al. (2021). ZFP451-mediated SUMOylation of SATB2 drives embryonic stem cell differentiation. Genes Dev. *35*, 1142–1160.

84. Ong, S.-E., and Mann, M. (2006). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat. Protoc. *1*, 2650–2660.

85. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

86. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526.

87. Simon, J.M., Giresi, P.G., Davis, I.J., and Lieb, J.D. (2013). A Detailed Protocol for Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE). Curr. Protoc. Mol. Biol. *102*, 21.26.1–21.26.15.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
H3	Abcam	Abcam Cat# ab1791; RRID:AB_302613
H3panAc	Millipore	Millipore Cat# 06-599; RRID:AB_2115283
H3K4me1	Abcam	Abcam Cat# ab8895; RRID:AB_306847
H3K4me2	Millipore	Millipore Cat# 07-030; RRID:AB_310342
H3K4me3	Abcam	Abcam Cat# ab8580; RRID:AB_306649
H3K9ac	Millipore	Millipore Cat# 07-352; RRID:AB_310544
H3K9me3	provided by T. Jenuwein	Serum
H3K27me3	Millipore	Millipore Cat# 17-622; RRID:AB_916347
H3K27Ac	Abcam	Abcam Cat# ab4729; RRID:AB_2118291
H4panAc	Millipore	Millipore Cat# 06-866; RRID:AB_310270
H4K16ac	Active Motif	Active Motif Cat# 39167; RRID:AB_2636968
OCT2	Santa Cruz	Santa Cruz Biotechnology Cat# sc-233; RRID:AB_2167205
RbBP5 (MLL subunit)	Bethyl	Bethyl Cat# A300-109A; RRID:AB_210551
YY1	Santa Cruz	Santa Cruz Biotechnology Cat# sc-1703, RRID:AB_2218501
IgG	Santa Cruz	Santa Cruz Biotechnology Cat# sc-2027; RRID:AB_737197
IgG	Santa Cruz	Santa Cruz Biotechnology Cat# sc-2025; RRID:AB_737182
IgG	Santa Cruz	Santa Cruz Biotechnology Cat# sc-2026; RRID:AB_737202
MBD2/3	Millipore	Millipore Cat# 07-199; RRID:AB_310423
EBF1 (6G6)	provided by E. Kremmer	Hybridoma supernatant
GAPDH	Calbiochem	CB1001
Actin	Sigma Aldrich	Sigma-Aldrich Cat# A2066; RRID:AB_476693
p300	Santa Cruz	Santa Cruz Biotechnology Cat# SC-585; RRID:AB_2231120
YY1	Santa Cruz	Santa Cruz Biotechnology Cat# sc-1703; RRID:AB_2218501
E2A	BD Biosciences	BD Biosciences Cat# 554077; RRID:AB_395228
SATB1	BD Biosciences	BD Biosciences Cat# 611182; RRID:AB_398716
SATB2	Santa Cruz	Santa Cruz Biotechnology Cat# sc-81376; RRID:AB_1129287
CHD8	Bethyl	Bethyl Cat# A301-224A; RRID:AB_890578
CTCF	Cell Signaling	Cell Signaling Technology Cat# 2899; RRID:AB_2086794
IRF4	Santa Cruz	Santa Cruz Biotechnology Cat# sc-377383
IRF8	Santa Cruz	Santa Cruz Biotechnology Cat# sc-6058; RRID:AB_649510
MZB1	provided by E. Kremmer	Hybridoma supernatant
THAP11	provided by E. Kremmer	Hybridoma supernatant
MOF	Bethyl	Bethyl Cat# A300-992A; RRID:AB_805802

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MSL1	Sigma Aldrich	Sigma-Aldrich Cat# SAB1306806
MSL1	provided by Asifa Akhtar	serum
MSL2	Sigma Aldrich	Sigma-Aldrich Cat# HPA003413; RRID:AB_1848659
POL2 Rpb1 NTD (D8L4Y)	Cell Signaling	Cell Signaling Technology Cat# 14958; RRID:AB_2687876
ERG	Abcam	Abcam Cat# ab92513; RRID:AB_2630401
CD3	BD Biosciences	BD Biosciences Cat# 612771; RRID:AB_2870100
CD19	BD Biosciences	BD Biosciences Cat# 552854, RRID:AB_394495
B220	BD Biosciences	BD Biosciences Cat# 553092, RRID:AB_398531
HSA (CD24)	BD Biosciences	clone M1/69
BP1	Biolegend	clone 6C3
CD43	BD Biosciences	BD Biosciences Cat# 553270, RRID:AB_394747
CD45.1	BD Biosciences	clone A20
CD45.2	BD Biosciences	clone 104
GR1	BD Biosciences	clone RB6-8C5
TCR-β	Biolegend	clone H57-597
IgM	Southern Biotech	clone 1020-09
IgM	BD Biosciences	BD Biosciences Cat# 553437, RRID:AB_394857
IgKappa	Invitrogen	clone MKAPPA04
IgLambda	Southern Biotech	clone 1060-03
IgLambda	BD Biosciences	clone R26-46
IgMa	BD Biosciences	clone DS-1
IgMb	BD Biosciences	clone AF6-78
Chemicals, peptides, and recombinant proteins		
dTAG13	Tocris	6605
Dynabeads^{TM} MyOne^{TM} Streptavidin C1	Thermo Scientific	65001
Deposited data		
proteomics μWT vs. μREV	This study	Table S1
proteomics μWT vs. μCoreREV	This study	Table S2
proteomics μWT vs. μE1^{MT1}	This study	Table S3
proteomics μWT vs. μΔMAR	This study	Table S4
Msl2^{KO} RNA-seq in *Slp65*^{KO} cell line	This study	Table S7
MSL2 ChIP in *Slp65*^{KO} cell line	This study	Table S6
YY1 ChIP in *Slp65*^{KO} cell line	This study	Table S6
H3K4 ChIP in *Rag1*^{KO} proB cells	Lin et al., 2010	Table S6
Experimental models: Cell lines		
38B9 pro B cell line	Alt et al., 1984[82]	N/A
Slp65^{KO} pre B cell line	Jumaa et al., 2001[42]	N/A
AMuLv transformed *Rag1*^{KO} pro B cell line	This study	N/A
Experimental models: Organisms/strains		
Mouse: *Mof* $^{+/-}$ C57BL/6	Pessoa-Rodrigues et al., 2020[41]	N/A
Oligonucleotides		
Primers for DNA baits	This study	Table S9

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primers for qRT-PCR	This study	Table S9
Primers for 3C assays	Guo et al., 2011[18]; Hao et al.,2015[49]	Table S9
gRNA sequences	This study	Table S9
Recombinant DNA		
pMYS-MSL2cDNA-IRES-GFP	This study	N/A
pCas9 (BB) 2A-GFP (pX458)	Addgene	48138
pAW63.YY1.FKBP.knock-in.BFP	Addgene	104371
Software and algorithms		
MaxQuant	https://maxquant.org/	N/A
Perseus	https://maxquant.net/perseus/	N/A
R package	https://www.r-project.org	N/A
Python	https://www.python.org	N/A
Picard	https://broadinstitute.github.io/picard/	N/A
Samtools	https://github.com/samtools/samtools	N/A
deepTools	https://deeptools.readthedocs.io/en/	N/A
Homer	http://homer.ucsd.edu/homer/	N/A
pheatmap	https://cran.r-project.org/web/packages/pheatmap/	N/A
Gviz	https://bioconductor.org/packages/release/bioc/html/Gviz.html	N/A
STAR-2.5.3a	https://github.com/alexdobin/STAR	N/A
trim_galore- 0.3.7	https://github.com/FelixKrueger/TrimGalore	N/A
featureCounts v2.0.0	https://subread.sourceforge.net/featureCounts.html	N/A
Limma v.3.56	https://bioconductor.org/packages/release/bioc/html/limma.html	N/A
Gencode.vM23	https://www.gencodegenes.org/mouse/release_M23.html	N/A
Reference genome mm10	https://genome.ucsc.edu	N/A
Fgsea	https://bioconductor.org/packages/release/bioc/html/fgsea.html	N/A
ImageJ software	https://quantitative-plant.org/software/imagej	Figure 2C

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Rudolf Grosschedl (grosschedl@ie-freiburg.mpg.de).

Material availability

The cell lines used in this work and commercially available reagents are indicated in the key resources table. The cell lines generated in this study are available without restriction from the lead contact.

Data and code availability

- Next Generation sequencing data generated in this study have been deposited under the accession number GEO: GSE242256 and are publicly available as of the date of publication.
- Proteomics data are available via ProteomeXchange with identifier ProteomeXchange: PXD045636, ProteomeXchange: PXD045637, ProteomeXchange: PXD045694 and ProteomeXchange: PXD045704.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

Wild type FVB mice and *Mof* $^{+/-}$ mice[41] were bred and maintained in the animal facility of the MPI under regular housing and specific pathogen-free (SPF) conditions. All animal procedures were performed according to protocols approved by the German authorities and Regierungspräsidium Freiburg. For the generation of transgenic embryos, 6–8 week old female FVB mice were mated with 8–12 week old FVB male mice, and fertilized eggs were obtained for microinjection of linearized plasmid DNA.[12] Microinjected eggs were implanted into FVB female foster mothers and fetal livers were obtained at E16.

All animal experiments with μE1^{mt3} mice were approved by the Institutional Animal Care and User Committees at NIH.

- Wild type FVB mice were used for the generation of transgenic fetal liver cells.
- *Mof* $^{+/-}$ mice were previously generated for a study of the role of *Mof* in early hematopoiesis.[41].
- μE1^{mt3} mice were generated at Johns Hopkins University, School of Medicine, Transgenic Mouse Core.
- *Rag2* $^{-/-}$ mice were purchased from Jackson Lab (stock no. 008449).

Cell lines

A-MuLV-transformed 38B9 and μ transgenic pro-B cell lines, as well as *Mof* $^{+/-}$ pro-B cells were used for the following purposes (a) analysis of the Eμ enhancer proteome; (b) ChIP analysis of histone modifications and MSL2 and YY1 DNA binding.

Primary cell cultures

Pre-B cells obtained from the bone marrow of *Slp65*^{KO} BALB/c mice[42] were a kind gift of Dr. P. Nielsen (MPI of Immunobiology and Epigenetics). Primary *Slp65*^{KO} pre-B cells were cultured and single cell-cloned in IMDM medium containing 10 ng/mL recombinant IL-7. *Slp65*^{KO} pre-B cells were used for (a) ChIP analysis of histone modifications and MSL2 and YY1 DNA binding; (b) CRISPR/Cas9-mediated gene editing; (c) ChIP-seq analysis and (d) RNA-seq analysis.

METHOD DETAILS

Generation of transgenic cell lines

pμ, pE1^{mt1}, pΔMAR and pμΔE1ΔMAR plasmid DNA, containing a functionally rearranged Igμ locus,[12] was linearized with Sal I and microinjected into fertilized eggs from female FVB mice that had mated with FVB males the previous night. At E16, mouse embryos were isolated and screened for the presence of injected DNA. Fetal liver cells from transgene-containing embryos were immortalized by infection with Abelson murine leukemia virus to generate pro-B cell lines.

CRISPR/Cas9 editing of *Slp65*^{KO} pre-B cells

Cells previously obtained from the bone marrow of *Slp65*^{KO} BALB/c mice[42] were subcloned to generate single cell clones. Cells from a single clone were electroporated with various gRNAs to generate μE1^{mt2}, *Msl2*^{KO} and YY1-FKBP knock-in cell lines. The CRISPR guide RNAs (gRNAs) were designed using the online tool from the Massachusetts Institute of Technology (http://crispr.mit.edu). The sequences of gRNAs are shown in Table S9. The gRNAs were cloned into pCas9 (BB) 2A-GFP (pX458; Addgene 48138). A minimum of three gRNAs were screened for each locus, and their efficiency was determined using the T7 endonuclease assay as described before.[83] The μE1^{mt2} cell line was generated by two gRNAs, positioned around the μE1 box. For the μE1^{mt2} cell line, CRISPR editing was exclusively accomplished on the VDJ-rearranged allele (Table S10). In all other cases both alleles were targeted by CRISPR gRNAs. The *Msl2*^{KO} cell line was generated by introducing a point mutation that induces a frameshift, eventually leading to the creation of a *de novo* stop codon (Table S10). The YY1-FKBP knock-in cell line was generated as described before.[47]

Generation of μE1^{mt3} pro-B cell lines

The μE1^{mt3} mutation was generated by replacing the μE1 box sequence in the germline *Igh* locus of C57BL/6 mice with a Tet operator sequence (TetO) by using CRISPR/Cas9-mediated gene editing. μE1^{mt3} mice were bred to *Rag2*$^{-/-}$ mice to generate *Rag2*^{KO} μE1^{mt3} mice. To generate *Rag2*^{KO} μE1^{mt3} pro-B cell lines, bone marrow cells from *Rag2*^{KO} μE1^{mt3} mice were immortalized by infection with Abelson murine leukemia virus. The sequences of the gRNA and its knock-in construct are shown in Table S9.

Cell cultures and SILAC labeling

A-MuLV-transformed cell lines were grown in RPMI 1640, supplemented with 10% FCS, 2 mM L-glutamine, penicillin and streptomycin. Primary *Slp65*^{KO} cell lines were grown in IMDM medium with 10 ng/mL recombinant IL-7, supplemented with 10% FCS, 2 mM L-glutamine, penicillin and streptomycin. Cells were incubated with 5%CO_2 at 37°C. Cells that were SILAC-labeled[84] were counted and washed one time with sterile 1xPBS. The number of cells in each culture were kept stable (less than 0.5x10^5 cells/ml) for 5 days to confirm that all cells are labeled. Then, the cells were resuspended into the according type of labeled medium. After 5 days, the cell

number can be increased and used as normal medium for any experimental purpose. RPMI medium without L-Arg and L-Lys was used because the labeled amino acids are lysine and arginine (K0R0, K6R4 and K8R10).

Intracellular FACS

10^6 cells were stained with viability dye (1:1000 dilution; cat. no. 65-0864-14) for 20 min. After washing, fixation and resuspension in permeabilization buffer, using the eBioscience Foxp3/Transcription Factor staining kit (cat. no. 00-5523-00), cells were incubated with FACS antibodies (key resources table) and used for flow cytometry.

Preparation of nuclei

Cells were grown to a volume of 3–4 L (ca. 1.0x10^6 cells/ml) and were harvested by centrifugation at 400 *g*, 4°C for 10 min. The supernatant was discarded and after 2 wash steps with PBS in 50 mL falcon tubes cells were pelleted at 400 g, 4°C for 10 min. The supernatant was discarded and the volume of sediment was recorded as "packed cellvolume: PCV". Then, the cells were resuspended in 5x PCV hypotonic HB buffer (10 mM NaCl, 20 mM Tris, pH 7.6, 3 mM $MgCl_2$, 1 mM DTT and protease inhibitor mix) and incubated on ice for a maximum of 10 min to let them swell. The cells were spun down at 400 g, 4°C for 5 min. The volume of the pellet was recorded as "swollen cell volume: SCV". After this step, the volume of the cells should have increased. The supernatant was discarded and the pellet was resuspended in 2x PCV HB buffer. The cells were transferred to a pre-cooled glass homogenizer on ice and homogenized with "loose" pistil (20 strokes). The cell membrane was disrupted by 5 strokes with "tight" pistil. An aliquot was taken prior to cell lysis and controlled by comparison to lysed cells (phase contrast light microscopy). The nuclei were sedimented by spinning at 1800 g, 4°C for 10 min. The supernatant was carefully taken off as cytosolic fraction and adjusted to 25% glycerin. The volume of cytosol and "nuclear pellet volume: NPV" were recorded. The nuclei were further resuspended in 1 volume NPV of HB buffer containing 25% glycerin. The nuclei were extracted by slowly adding 3M KCl to 450 mM KCl (final concentration) for at least 30 min on ice.[85] The nuclei were incubated for additional 10 min and were finally spun down in an ultracentrifuge at 125,100 g at 4°C for 30 min. Both cytosolic and nuclear fractions were snap frozen in liquid nitrogen. Before use, the nuclear extracts are dialyzed overnight at 4°C against dialysis buffer (20 mM HEPES, pH 7.9, 20% Glycerol, 110 mM NaCl, 0.2 mM EDTA).

qRT-PCR analysis of mRNA expression

RNA was extracted by NucleospinRNAII isolation kit (Macherey-Nagel) and subsequently reverse transcribed to cDNA using Oligo(dT) primer and SuperScriptII kit (Invitrogen). For quantitative RT-PCR, 1 μL cDNA of cell target was mixed with 3 μL of MilliQ water, 1μL of primer pair of gene of interest and 5 μL of SYBR Green PCR Master Mix (Promega) in a well of optical 96-well reaction plate (total volume 10 μL). The reaction was one replication of 50°C for 2 min, 2 replications of 95°C for 2 min and 45 replications of 95°C for 15 s and a final step at 60°C for 1 min. The additional dissociation graph is one replication of 95°C for 15 s, 60°C for 1 min and 95°C for 15 s. The analysis was done using Ct value approach at the exponential phase of PCR.

Co-immunoprecipitation (CoIP)

CoIP was done with either native nuclear extract or EGS cross-linked nuclear extract. For EGS crosslinking, EGS was added to a final concentration of 0.686 g/L. After 15 min at room temperature, the reactions were quenched with 100 mM glycine. The nuclear extracts were prepared from fixed cells using Dignam protocol. The antibodies were incubated with protein A Sepharose beads overnight and blocked for 1 h at 4°C using blocking buffer (20 mM HEPES, pH 7.9, 0.05 mg/mL, BSA, 0.05 mg/mL glycogen, 0.3 M KCl, 0.02% NP40, 2.5 mM DTT, 5 mg/mL polyvinylpyrrolidone). The reactions were incubated overnight at 4°C. The reactions were washed four times with buffer G (20 mM HEPES, pH 7.9, 10% glycerol, 100 mM NaCl, 1 mM EDTA, 10 mM potassium glutamate, 0.04% NP40, and protease inhibitor mix). The beads were eluted by Laemmli buffer. TCEP (a final concentration of 10 mM) and CAA (a final concentration of 40 mM) were added to the eluates. The reactions were incubated at 60°C for 30 min in the dark and warmed up to 70°C, 1 min before loading on the gel.

Chromatin immunoprecipitation (ChIP)

Cells (5x10^6) were fixed at room temperature with 1% (v/v) formaldehyde (5 min for histone, 15 min for transcription factors), and quenched with 100 mM glycine. Cells were washed three times with cold PBS and were lysed in 625 μL lysis buffer (50 mM Tris, pH 8.0, 10 mM EDTA and 1% (w/v) SDS) containing protease inhibitors. Chromatin was sheared by sonication to an average size of 500 base pairs and was then diluted 1:10 with dilution buffer (50 mM Tris, pH 8.0, 10 mM EDTA, 150 mM NaCl and 1% (v/v) Triton X-100). For histone ChIP, a 1.0-mL aliquot of the diluted chromatin was incubated first with antibodies overnight and next day with protein A Sepharose beads for 1 h at 4°C. For transcription factor ChIP, the antibodies were first bound to protein A Sepharose beads and blocked for 1 h at 4°C using blocking buffer (20 mM HEPES, pH 7.9, 0.05 mg/mL BSA, 0.05 mg/mL glycogen, 0.3 M KCl, 0.02% NP40, 2.5 mM DTT, 5 mg/mL polyvinylpyrrolidone). The diluted chromatin was added to the antibody bound protein A Sepharose beads. The reactions were incubated overnight at 4°C. The reactions were washed four times with wash buffer (20 mM Tris, pH 8.0, 2 mM EDTA, 0.1% (w/v) SDS, 1% (v/v) Triton X-100 and 150 mM NaCl) and then washed once with Tris-EDTA (20 mM Tris, pH 8.0, 2 mM EDTA). Chromatin was eluted twice with 50 μL elution buffer (1% (w/v) SDS, 0.1 M $NaHCO_3$) and was reverse-cross-linked by incubation of the chromatin immunoprecipitation samples overnight at 65°C. DNA was purified with the Qiaquick PCR purification kit (Qiagen) and was analyzed by quantitative PCR.

Native chromatin immunoprecipitation (native ChIP)

$2*10^7$ cells were grown and harvested to prepare the nuclei. The nuclei were resuspended in 400 μL MNase buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5 mM $CaCl_2$, 100 μg/mL BSA) and digested with 60 units of MNase (Thermo no.: EN0181) at 37°C shaking for 9 min (depends on cell types) to obtain mononucleosomes. The digestion was stopped by addition of EDTA and EGTA to 0.3 M and 0.25 M final concentration, respectively. Following a short spin down, the supernatant containing the DNA fragments was controlled by agarose gel electrophoresis. Then chromatin was subjected to antibody-protein A Sepharose immunoprecipitation as described above.

DNA affinity chromatography experiments

Biotinylated DNA baits generated by PCR were purified by phenol-chloroform extraction/DNA precipitation and immobilized to MyONE C1 streptavidin magnetic beads with the ratio 200 pmol DNA/100 μL beads. The beads were then kept in buffer DW (20 mM Tris-HCl, pH 8.0, 2 M NaCl, 0.5 mM EDTA, 0.01% NP-40). Before conducting the reverse ChIP experiment, the beads harboring the immobilized DNA baits are incubated for 1 h in blocking buffer (20 mM HEPES, pH 7.9, 0.05 mg/mL, BSA, 0.05 mg/mL glycogen, 0.3 M KCl, 0.02% NP40, 2.5 mM DTT, 5 mg/mL polyvinylpyrrolidone) at 4°C on a rotary wheel using 13.4 μL buffer per μL beads. Excess blocking buffer is removed by washing the beads with buffer G (20 mM HEPES, pH 7.9, 10% glycerol, 100 mM NaCl, 1 mM EDTA, 10 mM potassium glutamate, 0.04% NP40, and protease inhibitor mix). Dialyzed nuclear extracts were precleared by spinning down at 20,000 g for 10 min at 4°C and incubation with bulk MyONE C1 beads. Cleared nuclear extracts (~6–10 mg/mL) are adjusted to 10 mM potassium glutamate, 10 μg/mL poly (dI/dC) and 110 mM NaCl (final concentrations). The adjusted extracts were separately subjected to pre-blocked beads with respective DNA baits for 4 h at 4°C. After incubation, the beads were washed 4 times with buffer G. The beads were then eluted two times (for 5 μL beads carrying biotinylated DNA baits, 1st elution: 20 μL buffer G, 2 mM $MgCl_2$, 1 μL Benzonase, 2nd elution: 30 μL buffer G). The eluates were combined and adjusted to 1xLDS sample buffer/50mM DTT.

Proteomics

Samples derived from DNA affinity chromatography experiments were separate by 4–12% Bis-Tris NuPAGE gels and stained with colloidal Coomassie (InstantBlue). For each replicate, entire gel lanes were cut into evenly distant slices, which were processed by standard trypsin (Promega) in-gel digestion procedure. Following peptide clean-up by C18-STAGE tipping, tryptic peptides were analyzed by nanoLC-MS in DDA mode (Thermo Fisher Orbitrap XL + ETD or Q Exactive/Q Exactive Plus and Exploris 480 mass spectrometer) using a one column liquid-junction setup, in which the in-house packed (Dr. Maisch, Reprosil-Pur C18 AQ 120A beads) analytical capillary column (NewObjective, 360 μm OD, 75μm ID, 8 μm tapered open end) concomitantly served as the ESI (electrospray ionization) emitter. MS .raw data were processed by MaxQuant (SILAC data acquired on OrbitrapXL+ETD MS; label-free data acquired on Q Exactive and Exploris mass spectrometers) utilizing the Max LFQ algorithm[86] for label-free data and the output (proteingroups.txt) was further analyzed by standard R packages (quantile normalization of LOG2 transformed data in limma package, pheatmap.R) and the Perseus framework (Student's T test, Benjamini-Hochberg multiple testing correction, Volcano plots).

Formaldehyde-assisted isolation of regulatory elements (FAIRE)

FAIRE was performed as described[87] and analyzed with qPCR.

3C assay

As outlined previously,[18] $1x10^6$ to $2x10^6$ cells were lysed in 300 μL of lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.2% Igepal CA630 with protease inhibitors) and incubated for 20 min on ice. Cells were centrifuged 2500 *g* for 5 min at 4°C, and pellets were washed once in lysis buffer. Pellets were resuspended in 50 μL of 0.5% SDS and incubated for 10 min at 65°C. Water (145 μL) and 25 μL of 10% Triton X-100 were added to the samples and incubated for 15 min at 37°C. HindIII restriction enzyme (100 U) and 25 μL of NEB Cut smart buffer were added and incubation was extended (overnight at 37°C with shaking). The next day, the enzyme was inactivated for 20 min at 65°C. The ligation reaction was carried out overnight at 16°C by adding 120 μL of NEB T4 ligase buffer containing 10 mM ATP (NEB B0202), 100 μL of 10% Triton X-100, 3 μL of 50 mg/mL BSA, 720 μL of water, and 5 μL of T4 DNA ligase (NEB M0202). The day after, 50 μL of 20 mg/mL proteinase K and 120 μL of 10% SDS were added, and the samples were incubated overnight at 65°C. Last, 10 μL of 10 mg/mL RNase was added, and samples were incubated for 1h at 37°C. Following phenol chloroform purification, the DNA was precipitated using 1.6 vol. of 100% ethanol and 0.1 vol. of 3 M sodium acetate. After incubation for 1h at −80°C, samples were spun at 16000 rpm for 15 min at 4°C. Pellets were washed twice with 70% ethanol and dissolved in 100 μL of 10 mM Tris (pH 8.0). 3C ligation products were measured by quantitative PCR, and primers for the amplification of the "bait" sequence were used as an internal normalization control for each of the samples. The primers used for this study are listed in Table S9.

RNA-seq analysis

Total RNA was isolated using Qiagen micro RNA prep kit according to the manufacturer's instructions. rRNA-depleted total RNA was prepared using Ribozero magnetic beads for library preparation. The library was prepared with random hexamer primers. The 151bp paired-end reads were mapped to the mouse reference genome (mm10) using the STAR aligner (2.5.3a). Total read counts for the Gencode (vM23) reference genes were calculated using featureCounts (v2.0.0). The log-transformed total count per million (CPM) was calculated for each gene and genes with values below 0 were excluded. Differential gene expression was calculated by a linear

modeling approach with the use of functions available in the R package limma. For differential expression analysis, the functions lmfit, eBayes and contrast.fit available in limma were applied to the voom-transformed expression values. A set of significant differentially expressed genes was selected with an adjusted *p*-value cut-off of 0.05 after the Benjamini-Hochberg correction for multiple testing was applied. Gene set enrichment analysis was performed using fgsea. This was used to identify significantly enriched pathways in the datasets. A final set of genes involved in the selected pathways was generated by manual literature review. The heatmap was visualized using pheatmap package available in R environment.

ChIP-seq analysis

ChIP-seq reads from two replicates each of MSL2, YY1, and RNA polymerase II (POL2) ChIP were trimmed using trim galore (v 0.3.7) and mapped to the mouse reference genome (mm10) using bowtie2 (2.2.3) with default settings. Read duplicates and reads mapped to blacklisted regions were removed using picard, samtools and bedtools. The properly mapped reads from both replicates of MSL2 or YY1 ChIP and their corresponding input controls were used to identify peaks using MACS2 (v2.1.0) with a q-value cut-off of 0.05. The overlap of MSL2 and YY1 peaks was identified within 75 bp of the peak summit using bedtools. Analysis of *de novo* motifs within 75bp of the peak in MSL2-only, YY1-only and the common binding regions was performed using the findMotifsGenome program available in the homer package (v 4.7). The distribution of the ChIP signal around the peak clusters was visualized using deeptools (v.3.5.0). The ChIP signal normalized to 10 million reads (bedGraph format) was generated using the makeUCSCfile tool available in Homer (v.4.7) and further used to visualize the distribution of signals around selected loci using Gviz tools (v 1.44). The peaks are annotated for the associated genes using the annotatePeaks tool available in Homer (v 4.7) and used for overlapping with differentially expressed genes. In order to properly map the YY1, MSL2 and POL2 binding signals at the rearranged *Igh* locus allele in *Slp65*^{KO} pre-B cells, we obtained the BALB/c genome sequence (GCA_921997145.2_BALB_cJ_v3) from the UCSC genome database. The VDJ region identified by targeted sequencing/RNAseq was reconstructed *in silico* to reflect the rearranged *Igh* locus allele. The MSL2, YY1 and POL2 ChIP sequencing data were mapped to the reconstructed genome to visualize the binding pattern at the *Igh* locus. For visualization, the replicates were pooled and normalized to 10 million reads using homer. The dataset was visualized for the reconstructed region (OW971649.1: 106900000–106990000) using pyGenomeTracks tools.

Analysis of publicly available H3K4 methylation ChIP-seq data

The histone marks ChIP of *Rag1* $^{-/-}$ pro-B cells was obtained from publicly available datasets[52] (GSE21978). The aligned reads were converted to bedGraph and bigwig format using Homer. The MSL2-specific, common (MSL2-and YY1-shared) and YY1-specific peaks were converted to the mm8 version using the ucsc liftover tool. Histone marks within 3kb of the peak center were visualized using deep tools.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were obtained from independent biological replicates as described in the method details and indicated in the figure legends. Descriptions of data quantification and statistical analysis of experiments are provided in the respective methods sections or in figure legends. *p* values and the number of replicates (n) can be found in the figure legends.