To elucidate the functional role of Satb1 in T cell development and differentiation

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy

द्वारा / By
आयुष मधोक / Ayush Madhok

पंजीकरण सं. / Registration No.: 20163439

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor
प्रोफ़ेसर संजीव गलांडे / Prof. Sanjeev Galande

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE
2024

Dedicated to

“My Parents”

CERTIFICATE

Certifies that the work incorporated in the thesis entitled "**To elucidate the functional role of Satb1 in T cell development and differentiation**" submitted by **Ayush Madhok** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Date 16 April 2024

Prof Sanjeev Galande
(Supervisor)

Declaration by Student

Name of Student: Ayush Madhok

Reg. No.: 20163439
Thesis Supervisor(s): Prof. Sanjeev Galande

Department: Biology
Date of joining program: 01-08-2016

Date of Pre-Synopsis Seminar: 22-09-2023

Title of Thesis: To elucidate the functional role of Satb1 in T cell development and differentiation

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of Prof. Sanjeev Galande.

Date: 16 April 2024 Ayush Madhok

Acknowledgements

The journey of my PhD, and the completion of it was only possible due to the constant support of many people, and I'm grateful to them all. First and foremost, I would like to express my most sincere appreciation and thanks to my PhD advisor Prof. Sanjeev Galande. I express my gratitude to him for encouraging fresh thinking and setting up prompt partnerships to finish the several facets of my work. I sincerely appreciate all his help, helpful advice, timely criticism and understanding direction during my PhD. He has helped me overcome my pitfalls and grow as a researcher, and learn the nuances of becoming a good professional.

I express my gratitude to Dr. Girdhari Lal and Prof. Vineeta Bal, members of my thesis committee, for their insightful feedback and perceptive ideas provided during the yearly meetings. I would like to express special appreciation for Prof. Geeta Narlikar's invaluable assistance and critical ideas during a crucial time in my doctoral studies as well as for providing some reagents important for my work. I would like to thank Prof. Shubhada Chiplunkar for providing her support and expertise in conceptualizing and completing the work related to the γδ T cell project. I would like to thank Dr. Sajad Bhat for his excellent support in the γδ portion of my work. I would also like to acknowledge Prof. CS Yajnik and Dr. Satyajeet Khare for their valuable support and help for the blood-transcriptomics work. I would also like to thank Prof. Girish Deshpande for valuable discussions at a critical juncture.

I'm really thankful to IISER Pune for allowing me an opportunity to pursue my doctoral studies. I would like to thank the chair(s) of Biology department during my stay here; Prof LS Shashidhara, Prof. Sanjeev Galande, Dr. Thomas Pucadyil and Prof. Richa Rikhy for my unhindered work in the department. A special thanks to former directors; Prof. KN Ganesh and Prof. Jayant Udgaonkar, as well as the current director, Prof. Sunil S. Bhagwat for their support. I humbly thank the staff members of National Facility for Gene Function in Health and Disease, IISER Pune; Dr. Suraj Ingle, Dr. Sagar Tarate, Dr. Mahesh and Mr. Kalyan for their support in providing experimental animals during the course of my PhD. I would like to thank Ms. Anubha Pulwale for constantly helping me with breeding of multiple genotypes and providing the mice as per my experimental requirements. I thank Dr. Santosh Podder and Mr. Vijay Vittal and the IISER Pune Microscopy facility for their support. I am thankful to Mrinalini Virkar, Rupali Jadhav, Kalpesh Thakare, Shabnam Patil, Sandeep Bejjanki, Piyush Gadekar, Sandeep Sankpal, Mahesh Rote and all the members of the biology office for making this academic journey smooth, as well as Tushar Kurulkar and Sayalee Damle from the Academics department. I thank Mr. Sourabh Vaishnav and IISER Pune FACS facility for constant support and timely bookings which helped me complete my work. I would also like to thank Dr.

Siddhesh Kamat for his suggestions in MS sample preparations, and Mr. Saddam Shaikh for his support in processing samples and providing the MS data. I am grateful to Shiv Nadar University for providing support to facilitate many experiments critical for my work.

I wish to express my gratitude to every member of the SG lab for creating a welcoming environment to pursue my work. I would like to thank senior members during my early days for their support; Dr. Rini Shah and Dr. Manu Unni, as well as Dr. Ashwin Kelkar for his helpful critiques. I would like to give special appreciation to Dr. Indumathi Patta, with whom I worked with during my early years, for providing molecular biology training, continuous support, scientific exchanges, and inspiration to succeed in research. I thank all the students who worked with me and helped me in various projects and helped me develop experience in guidance, as well as leading projects; Isha, Devika, Prateek, Harshit, Rupal, Tejashree, Dhrisaj and Kunal. I would like to thank Isha Desai and Dr. Mohd. Tayyab for helping me with some of the experiments in phase separation and protein-related work. I would also like to thank Dr. Mukul Rawat and Dr. Abhishek for their help in troubleshooting during my learning phase.

The "Council of Scientific and Industrial Research" (CSIR), New Delhi, has my sincere gratitude for providing financial assistance for my doctoral studies. I thank Infosys and DST-SERB for their financial assistance, allowing me to visit and attend conference in Germany.

I had real good friends who were always there despite my unavailability at times; Gargi, Tayyab, Isha, Mukul, Indu and Shruti. They had seen and tolerated me during my ebbs and flows, especially Shruti, who is a dear friend.

I express my heartfelt gratitude to Prof. K. Kannan who not only taught me early on but inspired, grilled and pushed me in the direction of research. I'm grateful to Prof. Meenu Kapoor for her constant support and insightful suggestions during my dissertation, as well as Prof. Nimisha Sharma for her amazing teaching and sheer candidness to aspire toward.

Lastly, but most importantly, I'm indelibly thankful to my grandfather, my parents and my sister for their love, support, motivation and giving freedom to do what I enjoy. I have reached here surely because of their love, support and constant prayers.

Ayush Madhok

List of Abbreviations

ATAC	Assay for transposase accessible chromatin
Bcl2	B-cell leukemia/lymphoma 2
BCA	Bicinchoninic acid Assay
BCL11B	B-Cell CLL/Lymphoma 11B (Zinc Finger Protein)
bHLH	Basic helix-loop-helix
BSA	Bovine Serum Albumin
bp	base-pair
BURs	base unpairing regions
CD-	Clusters of Differentiation
CDS	Coding Sequence
CD8SP	CD8 Single Positive
CD4SP	CD4 Single Positive
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP-Sequencing
CLPs	Common lymphoid progenitors
CMPs	Common myeloid progenitors
Ct	threshold cycle
cTECs	Cortical Thymic Epithelial Cells
CTCF	CCCTC-Binding Factor
cDNA	Complementary DNA
CtBP1	C-terminal binding protein 1
DAPI	4',6-diamidino-2-phenylindole
DNase I	Deoxyribonuclease I
DAG	Diacylglycerol
3D	Three Dimentional
DN	Double Negative
DP	Double positive
DTT	dithiothreitol
EDTA	Ethylene di-amine tetra-acetic acid
EGTA	Ethylene glycol tetra-acetic acid
ETPs	Early thymic progenitors
ECL	Enhanced Chemiluminescence
E2A	Transcription Factor E2-Alpha
FPKM	Fragments Per Kilo base of transcript per Million mapped reads

FACS	Florescence Activated Cell Sorting
FBS	Fetal Bovine Serum
FOXP3	Forkhead box P3
GATA	Glutamyl-tRNA amidotransferase subunit-A
GATA3	GATA Binding Protein 3
GFP	Green Fluorescent Protein
GSK-3	glycogen synthase kinase-3
H3K27me3	Histone H3 lysine 27 trimethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K9me	Histone H3 lysine 9 trimethylation
HD	Homeodomain
HDAC	histone deacetylase
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HES1	Hairy and enhancer of split-1
Hi-C	Chromosome Conformation Capture with high-throughput sequencing
Hi-ChIP	Hi-C chromatin immunoprecipitation
HSCs	Hematopoietic Stem Cells
HP1	Heterochromatin Protein 1
HD	Homeo Domain
HRP	horseradish peroxidase
ITAM	Immuno-Receptor Tyrosine-based Activation Motif
IB	immunoblot
IFNγ	Interferon γ
IgH	immunoglobin heavy chain
IL	Interleukin
IP	immunoprecipitation
kb	Kilobase
KCl	Potassium Chloride
kDa	Kilo Dalton
Lck	Lymphocyte Cell-Specific Protein-Tyrosine Kinase
LEF1	Lymphoid Enhancer Binding Factor 1
M	Molar
MACS	Model-based Analysis of ChIP-Seq
Mb	Mega Base Pairs
mRNA	Messenger RNA
MgCl2	Magnesium Chloride

MAPK	Mitogen-activated protein kinase
MD	MAR binding domain
MAR	Matrix/Scaffold attachment regions
ESCs	embryonic stem cells
Nur77	Nuclear Receptor Subfamily 4 group A member 1
NP40	Nonidet P-40
NaCl	Sodium Chloride
NIPBL	nipped-B-like protein
MHC	Major histocompatibility complex
NLS	nuclear localization sequence
NuRD	Nucleosome remodeling and deacetylase complex
PBS	Phosphate Buffered Saline
PCAF	P300/CBP-associated factor
PCR	Polymerase Chain Reaction
PDZ	Post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1)
qRT-PCR	Quantitative Real-time Polymerase Chain Reaction
PVDF	Polyvinylidene Fluoride or Polyvinylidene Difluoride
PI3K	Phosphoinositide 3-kinase
PKC	protein kinase C
PMA	Phorbol 12-myristate 13-acetate
PML	Promyelocytic leukemia
PMSF	phenylmethylsulfonyl fluoride
PRC	polycomb repressive complex
RAG	Recombinase activating genes
RNA pol II	RNA Polymerase II
RNA-seq	RNA Sequencing
RIPA	Radio Immunoprecipitation Assay Buffer
RPMI-1640	Rosewell Park Memorial Institute-1640
Runx	Runt-related transcription factor
SATB1	Special AT-rich Sequence Binding Protein 1
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulphate- polyacrylamide gel electrophoresis
SMC	Structural Maintenance of Chromosomes
SP	Single positive
STAT	signal transducer and activation of transcription

SWI/SNF	SWItch/Sucrose Non-Fermentable
TADs	Topologically Associated Domains
TCF-1	Transcription factor-1
TCR	T cell receptor
TF	transcription factor
TGF-β	Transforming growth factor-β
Th	T-helper
Thpok	T-helper-inducing POZ/kruppel factor
Treg	T-regulatory cells
Tris-HCl	2-Amino-2-(hydroxymethyl) propane-1, 3-diol hydrochloride
TSS	Transcription start site
ULD	ubiquitin-like domain
Wnt	wingless type protein
Xist	X-inactive specific transcript
Yoyo	Oxazole Yellow (dimer)

Contents

Synopsis

As the early thymic precursors migrate from the bone marrow to the thymus, the T lineage commitment is initiated. Determined by the levels of surface receptors CD4 and CD8, the development of T cells is divided into a number of temporally and physiologically different phases(Germain, 2002). The precursors, which do not possess CD4 and CD8, are referred to as double negative (DNs). Upon migration from the cortical segments of the medulla, DN thymocytes evolve to become double positive (DP) cells (Rothenberg et al., 2008). DPs are chosen based on a selection procedure and have a full αβTCR on their cell surface. Thymocytes that have undergone positive selection go on to become CD4 or CD8 single positive thymocytes (Singer and Bosselut, 2004). These thymocytes after maturation go to the secondary organs such as spleen and lymph nodes so they may multiply and react to antigenic cues to build a strong immune response (Lancaster et al., 2018). A fundamental set of factors play vital functions in this sequence of events by effectively triggering or blocking particular genes in conjunction with external stimuli. Though they regulate distinct genes, several of these TFs also play a role in the development during the later T cell stages (Rothenberg, 2014; Yui and Rothenberg, 2014).

One such factor is SATB1. Special AT rich binding protein (SATB1) is a transcription factor and chromatin organizer that is enriched in T cells (Galande et al., 2007). Nuclear protein SATB1 expression occurs first in the DN4 stage and then increased in the DP and CD4 phases. It has been previously demonstrated that in naïve CD4 T lymphocytes (Gottimukkala et al., 2016; Stephen et al., 2017), TCR stimulation positively regulates SATB1. Additionally, it was shown that SATB1 in CD4 cells has a "bimodal" distribution (Gottimukkala et al., 2016). The need for several transcriptionally active units is decreased when SATB1 loops the MHC class-I locus into a transcriptionally active area (Kumar et al., 2007). It has also been demonstrated that Th2 specific cytokine genes are regulated transcriptionally and looped (Cai et al., 2006).

Recent research using SATB1 null mice (Alvarez et al., 2000) as well as conditional knockout (cKO) has demonstrated that SATB1 deficiency partially affects positive selection (Kondo et al., 2016). Furthermore, it has been demonstrated that SATB1 is necessary for both the de-repression of the CD8a gene and ThPOK expression (Kakugawa et al., 2017). Moreover, it has also been shown that SATB1 directly binds to the promoter region of PD1 to control its expression (Stephen et al., 2017). The molecular processes behind SATB1's involvement in thymocyte formation and naïve T cell activation remain unclear despite a plethora of investigations.

My doctoral study focuses on the importance of SATB1 in maintaining the gene regulation during T cell development via TCR activation by maintaining appropriate 3D chromatin architecture. The work done during the period is divided into four chapters, summarized below:

1. **Satb1 regulates genome conformation and gene expression along with the cohesin complex in T cells**

In this study, we demonstrate that the T lineage-enriched chromatin organizer Satb1 shares majority of its binding sites across the genome in conjunction with the Cohesin complex as well as Ctcf in DP thymocytes. Genomic analyses revealed that Satb1 and Smc1a share multiple features such as binding motif, genomic locations and TSS binding. We confirm that Satb1 and Smc1a interact in vivo as well as in a purified setup. Satb1 shows strong interaction with the C-terminal regions of Smc1a, while Smc1a shows a preference for the ULD/PDZ domain of Satb1.

Moreover, Satb1 and Smc1a synergistically regulate the expression of CD3, a key surface receptor on T cells (Gascoigne et al., 2011), via directly binding its promoter. In the periphery, we observe that $CD4^+$ T cells exhibit dysregulated activation phenotype in Satb1 and Smc1a KO cells, indicating their collaborative effect on TCR signaling. Collectively, we demonstrate a molecular mechanism by which Satb1 mediates chromatin looping via its dynamic DNA binding and recruitment of the Cohesin complex to orchestrate 3D genome organization during T cell differentiation.

2. **Identification and characterization of liquid-like condensate formation by Satb1 in thymocytes**

We describe the ability of Satb1 to undergo condensate formation in this chapter. First, we determine the level of atypical characteristics in Satb1's amino acid composition. Next, we evaluated the biophysical characteristics of phase separation in vitro, including dependency on concentration, nucleic acid binding, and fluidity. Additionally, we investigated the particular domains implicated in Satb1's demixing behavior. We demonstrate that Satb1 produces nuclear foci in several phases of thymocyte development, which is consistent with phase behavior. In DP thymocytes, we also examined the dynamic nature of these foci. We show that nuclear foci of Satb1 co-localize with those of Smc1a, which also exhibit foci-like patterning. In contrast, Ctcf exhibits a more dispersed localization throughout different phases of T cell development. Finally, we demonstrate that the phase shifts of Satb1 in vitro is significantly influenced by certain amino acid residues of Satb1 linked to neurological disorders

and cancer. The capacity of Satb1 to undergo condensate formation may collectively increase its role in the control of transcription and genomic organization.

3. Role of Satb1 in regulating TCR-dependent genes in conjunction with NFAT family proteins during T cell development

Development of thymocytes is highly dependent on TCR signaling, which receive the signal in a spatiotemporally controlled manner (Love et al., 2000). SATB1, a T-lineage enriched chromatin organizer, controls the expression of various genes important in thymocyte development, and it is under the control of TCR signal (Gottimukkala et al., 2016; Patta et al., 2020). SATB1-deficient thymocytes do not develop beyond double positive (DP) stage and SATB1 knockout (KO) $CD4^+$ T cells undergo apoptosis upon TCR activation (Kondo et al., 2016), underscoring the crucial role of SATB1 in thymic selection. The mechanisms by which SATB1 mediates its regulatory function in developing thymocytes is only partly understood. We observed that in developing as well as naïve T-cells, expression of Satb1 is contingent on the strength of TCR stimuli in a linear manner. Furthermore, upon TCR activation of thymocytes, SATB1 interacts with nuclear factor of activated T-cells (NFAT), which we observe to be important for regulation of the transcriptional activity of numerous T cell specific genes such as Irf4, Dusp22 and IL2. Additionally, SATB1 has reduced occupancy on many of the target genes –which we observe to be enriched upon T cell activation –if NFAT is inhibited. We therefore conclude that these two proteins cooperate to regulate T cell specific key genes upon TCR mediated activation.

4. Dynamic modulation in the transcriptome of γδ T cells upon T cell activation and cross-talks with Notch signaling

Gamma delta (γδ) T cells, especially the Vγ9Vδ2 subtype have been implicated in cancer therapy and thus have earned the spotlight in the past decade (Zhao et al., 2018). Although one of the most important properties of γδ T cells is their activation by phosphoantigens which are, intermediates of the Mevalonate and Rohmer pathway of isoprenoid biosynthesis (Correia et al., 2009; Rigau et al., 2020), such as IPP and HDMAPP respectively, the global effects of such treatments on Vγ9Vδ2 T cells remain elusive. Here, we used the high-throughput transcriptomics approach to elucidate the transcriptional changes in human Vγ9Vδ2 T cells upon HDMAPP, IPP and anti-CD3 treatments in combination with IL2 cytokine stimulation. These activation treatments exhibited a dramatic surge in transcription with distinctly enriched pathways. We further assessed the transcriptional dynamics upon inhibition of Notch signaling coupled with activation treatments. We observed that the metabolic processes are most

affected upon Notch inhibition via GSI-X. The key effector genes involved in gamma-delta cytotoxic function were downregulated upon Notch blockade even in combination with activation treatment, thus, showing a transcriptional interplay between TCR signaling and Notch signaling in Vγ9Vδ2 T cells. Upon Notch inhibition, genes regulated by SATB1 upon γδ T cell activation via IPP, are dysregulated, suggesting its role at the juncture of both the pathways. Collectively, we show how activation of TCR signaling by phosphoantigens or anti-CD3 changes the transcriptional status of Vγ9Vδ2 T cells along with IL2 stimulation, and how blockade of Notch signaling can affect this activation.

References

Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev *14*, 521-535.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.

Correia, D.V., d'Orey, F., Cardoso, B.A., Lanca, T., Grosso, A.R., deBarros, A., Martins, L.R., Barata, J.T., and Silva-Santos, B. (2009). Highly active microbial phosphoantigen induces rapid yet sustained MEK/Erk- and PI-3K/Akt-mediated signal transduction in anti-tumor human gammadelta T-cells. PLoS One *4*, e5657.

Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev *17*, 408-414.

Gascoigne, N.R., Casas, J., Brzostek, J., and Rybakin, V. (2011). Initiation of TCR phosphorylation and signal transduction. Front Immunol *2*, 72.

Germain, R.N. (2002). T-cell development and the CD4-CD8 lineage decision. Nat Rev Immunol *2*, 309-322.

Gottimukkala, K.P., Jangid, R., Patta, I., Sultana, D.A., Sharma, A., Misra-Sen, J., and Galande, S. (2016). Regulation of SATB1 during thymocyte development by TCR signaling. Mol Immunol *77*, 34-43.

Kakugawa, K., Kojo, S., Tanaka, H., Seo, W., Endo, T.A., Kitagawa, Y., Muroi, S., Tenno, M., Yasmin, N., Kohwi, Y.*, et al.* (2017). Essential Roles of SATB1 in Specifying T Lymphocyte Subsets. Cell Rep *19*, 1176-1188.

Kondo, M., Tanaka, Y., Kuwabara, T., Naito, T., Kohwi-Shigematsu, T., and Watanabe, A. (2016). SATB1 Plays a Critical Role in Establishment of Immune Tolerance. J Immunol *196*, 563-572.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galande, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol *9*, 45-56.

Lancaster, J.N., Li, Y., and Ehrlich, L.I.R. (2018). Chemokine-Mediated Choreography of Thymocyte Development and Selection. Trends Immunol *39*, 86-98.

Love, P.E., Lee, J., and Shores, E.W. (2000). Critical relationship between TCR signaling potential and TCR affinity during thymocyte selection. J Immunol *165*, 3080-3087.

Patta, I., Madhok, A., Khare, S., Gottimukkala, K.P., Verma, A., Giri, S., Dandewad, V., Seshadri, V., Lal, G., Misra-Sen, J.*, et al.* (2020). Dynamic regulation of chromatin organizer SATB1 via TCR-induced alternative promoter switch during T-cell development. Nucleic Acids Res *48*, 5873-5890.

Rigau, M., Ostrouska, S., Fulford, T.S., Johnson, D.N., Woods, K., Ruan, Z., McWilliam, H.E.G., Hudson, C., Tutuka, C., Wheatley, A.K.*, et al.* (2020). Butyrophilin 2A1 is essential for phosphoantigen reactivity by gammadelta T cells. Science *367*.

Rothenberg, E.V. (2014). Transcriptional control of early T and B cell developmental choices. Annu Rev Immunol *32*, 283-321.

Rothenberg, E.V., Moore, J.E., and Yui, M.A. (2008). Launching the T-cell-lineage developmental programme. Nat Rev Immunol *8*, 9-21.

Singer, A., and Bosselut, R. (2004). CD4/CD8 coreceptors in thymocyte development, selection, and lineage commitment: analysis of the CD4/CD8 lineage decision. Adv Immunol *83*, 91-131.

Stephen, T.L., Payne, K.K., Chaurio, R.A., Allegrezza, M.J., Zhu, H., Perez-Sanz, J., Perales-Puchalt, A., Nguyen, J.M., Vara-Ailor, A.E., Eruslanov, E.B.*, et al.* (2017). SATB1 Expression Governs Epigenetic Repression of PD-1 in Tumor-Reactive T Cells. Immunity *46*, 51-64.

Yui, M.A., and Rothenberg, E.V. (2014). Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol *14*, 529-545.

Zhao, Y., Niu, C., and Cui, J. (2018). Gamma-delta (gammadelta) T cells: friend or foe in cancer development? J Transl Med *16*, 3.

Chapter 1: Satb1 regulates genome conformation and gene expression along with the cohesin complex in T cell development

1.1 Introduction

T lymphocytes are one of the two major arms of the adaptive immune system, the other being B cells, are responsible for mediating -and helping in- cell-based immune responses to prevent the manifestation of numerous diseases. T cells develop and mature in a highly specialized environment -an organ called thymus- imbued with complex and spatio-temporally restricted signaling processes. Upon screening via TCR signaling, only the correctly rearranged, moderately self-reactive T cells undergo maturation and emigration from the thymus into the secondary lymphoid organs to execute their distinct functions. After maturing, the thymocytes migrate to secondary organs like the lymph nodes and spleen, where they can proliferate and respond to antigenic stimuli to develop a potent immune response. In this series of events, a basic set of factors work together to efficiently trigger or inhibit certain genes in response to external inputs. Even though they control different genes, a few of these TFs are also involved in the maturation of subsequent stages of T cells.

1.1.1 T cell development – stages

In the thymus, T cells progress through various stages typically demarcated by the expression of co-receptors CD4 and CD8 on their surface, starting from double negative (DN), progressing to double positive (DP) followed by single positive (SP, $CD4^{+}$ or $CD8^{+)}$ stage (Germain, 2002). Thymocytes, depending on the stage of development, require migration to distinct thymic compartments, namely cortex and medulla (Lancaster et al., 2018), where they encounter different populations of thymic epithelial cells (TECs) (Abramson and Anderson, 2017). Distinct TECs from cortex and medulla provide unique signals to these migrating thymocytes, providing cues to undergo positive and negative selection, as well as the guiding chemokines for thymocyte migration (Lancaster et al., 2018). Given the great variability in the MHC alleles and the stochastic nature of TCR production, positive selection is essential for later T cell development (Carpenter and Bosselut, 2010). The early commitment, TCR dependent selection process as well as the key factors involved in mediating the timely expression of stage specific genes are discussed below.

1.1.2 Early development

T cell development initiates when the bone-marrow (BM) derived multipotent progenitors or Common Lymphoid precursors (CLP), which also give rise to B cells, enter the thymus through the cortical-medullary junction (CMJ) (Zhang and Bhandoola, 2014). After immigration, the early thymic progenitors (ETPs) in the cortex encounter Notch ligands which seals their commitment to a T cell fate (Hosokawa and Rothenberg, 2021). The DN stage is sub-divided into 4 distinct stages (DN1 through DN4) characterized by the surface expression of CD44 and CD25 (Rothenberg et al., 2008; Yui and Rothenberg, 2014). Receiving the Notch signal triggers the expression of CD25, progressing ETP/DN1 toward the DN2a stage ($CD25^{+}CD44^{+}$) (Rothenberg et al., 2008; Yang et al., 2010).

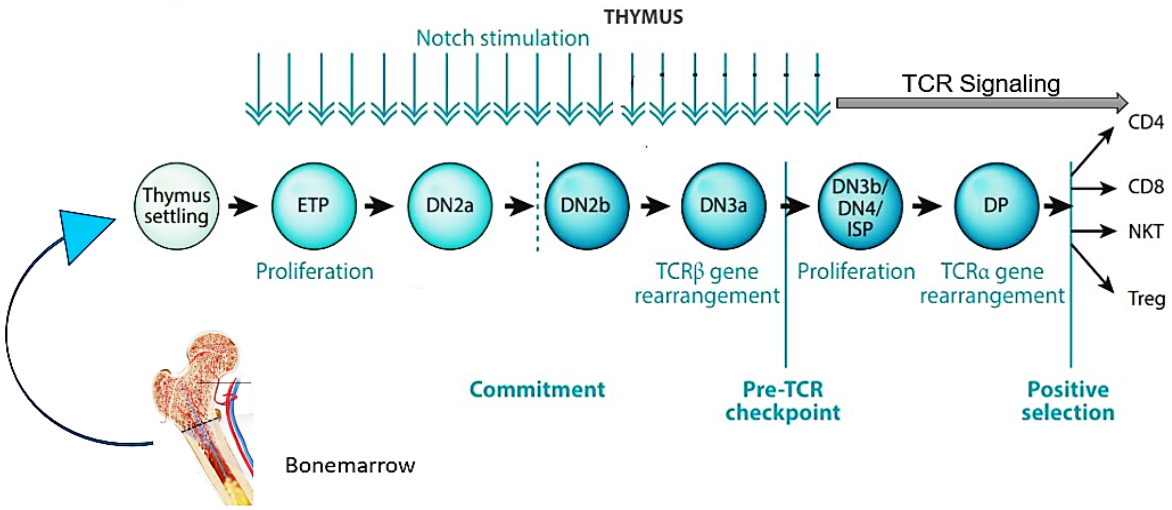


Figure 1.1.1: A schematic view of early and late phases of T cell development. Early T cell developmental stages initiate when the common lymphoid progenitor (CLP) which still retains the potential to undergo different lymphoid fates, migrates from bone marrow (BM) to the thymus and becomes early thymic precursors (ETPs). Here, upon receiving Notch signaling, they undergo T cell commitment and proceed through various stages of double negative (DN, devoid of CD4 and CD8 co-receptors) cells. After TCRβ rearrangement, the thymocytes prepare to release notch dependency and gain TCR responsiveness. DP thymocytes with functional TCRαβ expression undergo positive and negative selection upon interaction with self-peptide MHC I or II to undergo differentiation process to form multiple subsets such as CD4, CD8, Tregs and NKT cells, which then migrate to the periphery. Adapted from Rothenberg et al., 2014.

After the pre-commitment stage, during DN2b and DN3a ($CD25^{+}CD44^{-}$) stages, as RAG proteins are highly expressed (recombination of the T cell receptor genes (TCR) β, γ and δ occur to generate unique TCR heterodimers (Krangel, 2009; Yui and Rothenberg, 2014). Due to the random nature of recombination, a small population of thymocytes generating γδ rearrangement successfully results in selection of the γδ T cells (Dutta et al., 2021; Yui and Rothenberg, 2014). DN3b cells with TCR β chain rearrangement undergo β-selection with a functional pre-TCR complex consisting of CD3 proteins and an invariant TCRα chain, pre-Tα (Dutta et al., 2021; Nitta et al., 2008), selecting only the cells with responsive TCR β chains and progressing to DN4 stage ($CD25^{-}CD44^{-}$), throughout which they lose Notch dependence (Rothenberg, 2011; Yui and Rothenberg, 2014) and undergo proliferation (Kurd and Robey, 2016). Thymocytes that proliferate upregulate CD4 and CD8 co-receptors forming DP thymocytes. During DP stage, rearrangement of the *TCR α* locus occurs, which replaces the invariant T-α chain and functional TCR αβ receptor expresses on the DP thymocyte (Figure 1.1.1) (Kearse et al., 1995; Starr et al., 2003; Sun et al., 2023; Yui and Rothenberg, 2014).

1.1.3 Late development -Thymic selection and lineage commitment by TCR signaling

Once a functional TCR is expressed on the DP cell surface, all further development is contingent on the specificity (Figure 1.1.2A), the duration as well as the strength of the TCR signal that ensues from the interaction with the self-ligands expressed by the thymic epithelium (TECs) (Klein et al., 2014; Love et al., 2000; Singer et al., 2008; Starr et al., 2003)

1.1.3.1 Positive selection

One of the most critical steps that occur upon receiving TCR signal in the thymic cortex is positive selection. It is initiated when the TCR on a DP thymocyte forms a complex with any MHC molecules presented by the cTECs (Starr et al., 2003; Zerrahn et al., 1997). This acts as a survival step to only enrich DP clones (~5%) whose TCR can recognize antigens with the MHC complex, thus being the 'MHC-restricted' thymocytes (Zerrahn et al., 1997). There is a 3-4 days window during which the DP thymocytes can undergo positive selection, initiated by an interaction with self-pMHC, failing which triggers apoptosis (Alam et al., 1996; Takahama, 2006), referred here as 'death by neglect' (Figure 1.1.2B). Individual thymocytes also get multiple opportunities during this time-frame to edit out non-productive variable chains of TCRα and recombine different variable (V) and joining (J) on the *TCR α* locus, producing a new TCR αβ receptor, via a process called 'processive rearrangement' (Hoffman et al., 1996; Krangel, 2009).

1.1.3.2 Lineage commitment

Concomitant with positive selection, thymocytes selected for self-pMHC recognition are also tuned to their MHC reactivity. For instance, MHC-I reactivity of TCR on the DPs is different from that of MHC-II presented by the cTECs. DP thymocytes that recognize MHC-II develop into $CD4^+$ SP thymocytes, losing CD8 surface expression. While those which recognize MHC-I develop into $CD8^+$ thymocytes and lose their CD4 expression. CD4 and CD8 are of primary importance, since their binding of MHC-II or MHC-I are critical in

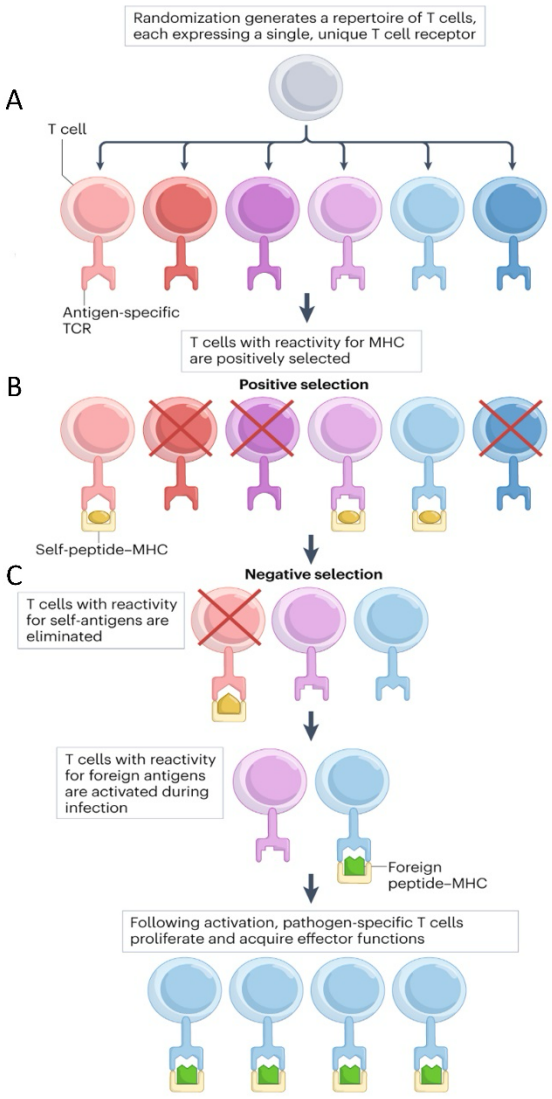


Figure 1.1.2: Clonal selection mechanism during T cell development. **A**. As the T cells become double positive, they express a unique TCR αβ on their surface, which is dependent on how the V(D)J recombination took place in that cell. **B**. Via the functional TCR αβ

expressing DP cells interact with the pMHC presented by the thymic epithelium with different affinity. The T cells which display a reactivity to the pMHC are positively selected and are migrated to medulla. The cells which do not show any reactivity are clonally deleted **C**. In the medulla, positively selected T cells which display high affinity toward pMHC presented by the mTECs undergo negative selection. Those with low-moderate affinity survive and mature. **D**. Among different mature T cells (with various combinations of TCR αβ) in the secondary lymphoid organs, the ones which recognize a foreign antigen via pMHC presentation during an infection gets clonally expanded, thus mounting an immune response. Reproduced from Ashby and Hogquist, 2023.

the process (Adoro et al., 2008; Robey et al., 1991; Shinzawa et al., 2022; Zeidan et al., 2019). There are three schools of thought on how the commitment takes place – stochastic, instructive and kinetic models that explain the initiation of lineage choice upon receiving the TCR signal. Stochastic model posits that the CD4/CD8 lineage commitment is a random event (Davis et al., 1993), suggesting that upon successful engagement of correct co-receptor with its cognate MHC complex – CD4 with MHC-II and CD8 with MHC-I – the thymocytes receive survival signal and downregulates the other co-receptor. The instructive model suggests the TCR signal strength drives the lineage choice (Iwashima, 2003; Singer, 2002; Starr et al., 2003). TCR signal intensity is much higher in a CD4:pMHC-II interaction than CD8:pMHC-I interaction. The model suggests that the higher strength of TCR signal produced by CD4:MHC-II interaction drives lineage choice toward CD4 SP leading to CD8 downregulation (Iwashima, 2003). Whereas relatively weaker signal by CD8-pMHC-I leads to CD8 SP commitment, downregulating CD4 expression. Although tracking lineage committed thymocytes with Nur77 GFP mice suggests that a direct correlation could not be established between the TCR signal intensity and the lineage choice (Malhotra et al., 2016).The kinetic signaling model somewhat explains exceptions in the strength model. It suggests that TCR signaling during positive selection induces downregulation of CD8, producing an intermediate $CD4^{+}CD8^{lo}$ population which are still bi-potential for commitment to either lineage (Singer, 2002). Further, an uninterrupted TCR signal promotes differentiation toward CD4 lineage, whereas a cessation in the signal directs the intermediate population toward CD8 SP lineage commitment.

1.1.3.3 Negative selection

Negative selection is the process of selectively ‘not selecting’ the cells with high-affinity interaction of their TCR and self-MHC. The high-affinity interactions trigger programmed cell

death -apoptosis (Pobezinsky et al., 2012; Stritesky et al., 2012) in these thymocytes (Figure 1.1.2C). Both cortex residing DP thymocytes and newly lineage committed CD4/CD8 SP, which migrated to medulla upon positive selection, are highly sensitive to negative selection or clonal deletion. Negative selection is distinct from positive selection as it is triggered only by high-affinity TCR:MHC interaction, whereas positive selection is their favored by low-affinity interactions (Figure 1.1.3). One of the key regulators of clonal deletion is BIM, a pro-apoptotic protein (Bouillet et al., 2002).

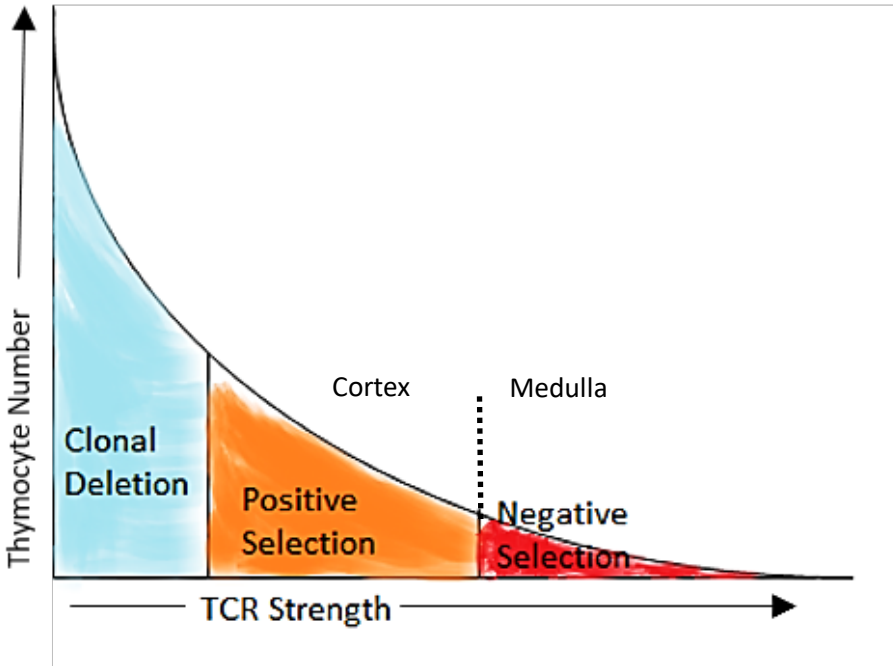


Figure 1.1.3: **TCR strength drives the selection process**. A majority of DP thymocytes with non-functional TCR αβ fail to bind self-pMHC in the thymic cortex, thus undergo apoptosis, a form of programmed cell-death. Positively selected thymocytes undergo another round of self-pMHC binding in the thymic medulla. The T cells which bind the pMHC with high affinity are subjected to apoptosis, hence negatively selected. A few of those escape negative selection and undergo a lineage commitment into regulatory T cells (Tregs).

1.1.4 Signaling mechanisms during T cell development

In the thymic microenvironment, various intracellular signaling cascades play synergistic and antagonistic roles during different stages of T cell development. But the most important exogenous signals from thymic epithelium are Notch and TCR signaling, occurring in early and late T cell developmental stages, respectively, mediate T cell commitment and lineage fate choice.

1.1.4.1 Notch signaling

Notch signaling is one of the most prominent intracellular pathways that mediates early T cell development and commitment, where cTECs provide Notch ligands to the progenitor populations entering the thymus. Notch1 deletion or HSPC-specific conditional knockout of Notch ligand delta-like 4 (DLL4) demonstrated a complete T lineage developmental blockade and occurrence of B cells in the thymus (Hozumi et al., 2008; Koch et al., 2008; Radtke et al., 1999). Whereas introduction of Notch signal to bone-marrow derived non-T cell progenitors induce a T cell developmental regime, both ex -and in vivo (Hozumi et al., 2003; Schmitt and Zuniga-Pflucker, 2002). Mechanistically, upon ligation of the notch receptor on the cell-surface and the notch ligand presented by cTECs, enables proteolytic cleavage of Notch and its intracellular domain (ICN) is released. ICN undergoes nuclear translocation and cooperates as a co-activator for recombination signal binding protein for immunoglobulin kappa J region (RBPJ), and activates cognate Notch-targets such as TCF1 and GATA3, and later BCL11B. Notch signaling provides survival and proliferation signal to the pro-T cells in phase 1 and phase 2 of early development as well as promote passage through the β selection stage (Figure 1.1.4) (Ciofani and Zuniga-Pflucker, 2006; Wolfer et al., 2002; Yashiro-Ohtani et al., 2010).

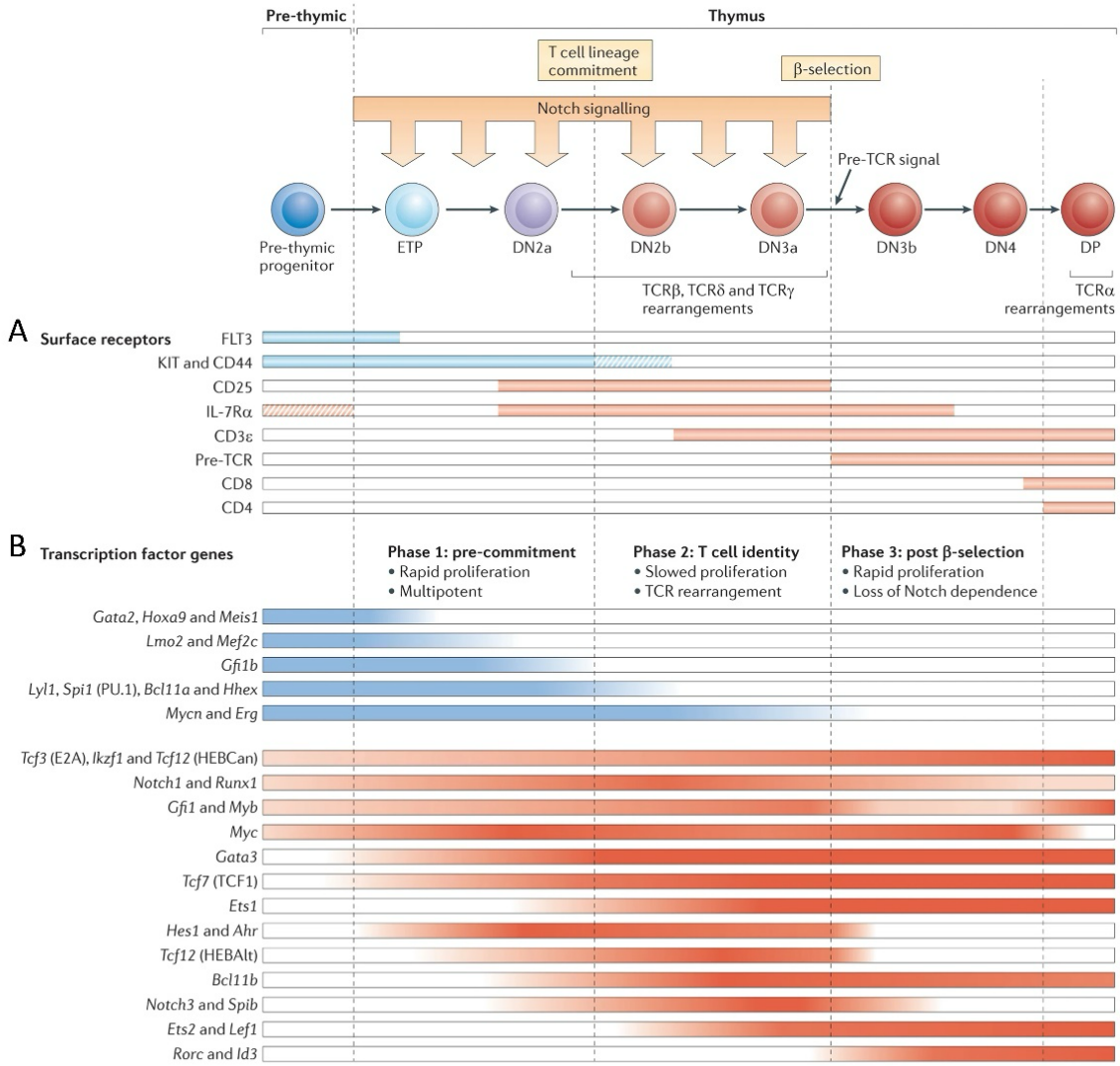


Figure 1.1.4: Notch signaling mediates early T cell development. The formation of adult mouse T cells from lymphoid precursors in the thymus in the presence of Notch ligands. **A.** The phases along with receptors that are utilized in these phases of T cell maturation. **B.** Important transcription factor mediated gene expression patterns in different stages are displayed. The variations in gene expression are depicted as the stages progress. Reproduced from Yui and Rothenberg, 2014.

1.1.4.2 TCR signal

As the pro-T cells progress from DN4 to DP stage, the MHC-restricted and β selected cells undergo a complete rearrangement of *TCRα* locus to form a functional TCR αβ receptor, along with expression of co-receptors CD4 and CD8. Upon their engagement with the pMHC-I or pMHC-II complex expressed on the TECs, conformational changes and receptor aggregation

occurs at the 'immunological synapse' (Cantrell, 1996; Smith-Garvin et al., 2009). This triggers activation of lymphocyte specific protein tyrosine kinase (PTK) such as Lck, Zap70 and Fyn (Samelson, 2002), which in turn phosphorylates various tyrosine residues on the cytoplasmic tails of the receptor protein CD3 (Guirado et al., 2002; Hwang et al., 2020; Molina et al., 1992; van Oers et al., 1996). Specifically, a unique motif presents on the cytoplasmic chains of CD3 proteins called immunoreceptor tyrosine-based activation motif (ITAM) gets phosphorylated by the PTKs Lck and Fyn (Gascoigne et al., 2011; Iwashima et al., 1994; Palacios and Weiss, 2004). Phosphorylation of ITAMs allows loading of Zap70 which activates downstream signaling cascades by bringing adapter proteins and additional signal transducers (Neumeister et al., 1995; Wang et al., 2010), to initiate the intracellular events of the TCR signal (Figure 1.1.5).

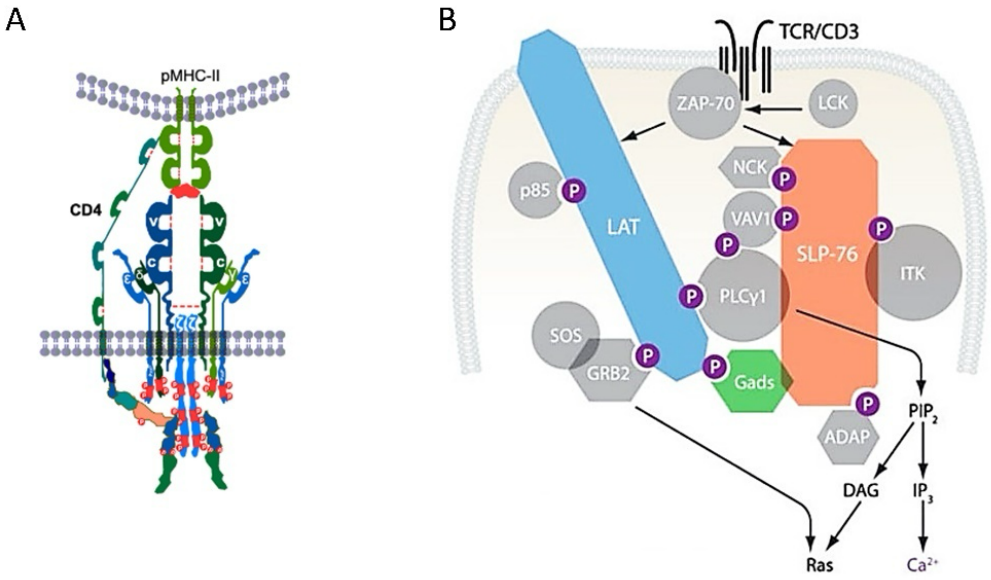


Figure 1.1.5: TCR activation- the proximal signaling. A. Both CD3ζ and CD3ε remain membrane-bound in resting T cells. Whereas CD3ζ gets released into the cytoplasm upon ITAM phosphorylation by LCK, thus producing a docking site for ZAP70. ZAP70 gets phosphorylated which then enables tyrosine phosphorylation on additional ITAM residues. **B.** The current knowledge of mechanism by which TCR activation connects to intracellular pathways such as Ca2+ influx and activation of Ras. The adaptor proteins LAT, Gads, and SLP-76 establish a multimolecular signaling complex that connects PTKs to downstream pathways. In order to phosphorylate tyrosine residues on LAT, ZAP70 is activated by Lck. This causes LAT to attract Gads and SLP-76. Multiple SH2-harboring effector molecules and adaptor proteins are recruited when ZAP70 phosphorylates SLP-76. Additionally connecting effectors to adapters and aiding in the complex's stability are SH3 domains, or the overlapping regions. Reproduced from Shah et al., 2021 (A) and Smith-Garvin et al., 2009 (B).

1.1.5 Transcription factors orchestrating developmental phases via chromatin states

In some myeloid lineages, distinctive sets of transcription factors are uniquely expressed; for instance, EBF1 and PAX5 in B cell development (Lin et al., 2010; Schebesta et al., 2007). However, there is no unique 'master-TFs' for T cell identity. Almost all the factors that promote T cell lineage can function to determine other lineages as well. For T cell identity, in fact it is the combined effect of overlapping, yet controlled expression of these factors along with Notch signal that drives T cell fate. Similarly, the TCR-dependent differentiation also rely on expression of multiple factors and their interplay with the chromatin states.

The indispensable factors for T lineage commitment are TCF1, GATA3, E2A, BCL11B, RUNX and IKAROS family members (Rothenberg et al., 2008; Seo and Taniuchi, 2016; Thompson and Zuniga-Pflucker, 2011; Yang et al., 2010; Yui and Rothenberg, 2014). Most of these factors' expression is limited to early development whereas some retain their expression in the later TCR dependent stages. The chromatin state shifts dramatically from the ETP stage to the TCR rearranged DP stage, including topologically interacting DNA, accessibility of chromatin, DNA methylation and histone modifications (Hu et al., 2018; Ji et al., 2010; Yoshida et al., 2019; Zhang et al., 2012). The specific combination of TFs with signals from thymic epithelium coordinate the switches driving the T cell fate and lineage choice (Figure 1.1.6).

1.1.5.1 TCF1

As early as ETP (DN1) stage, Tcf7 (encoding Tcf1, the high mobility group box transcription factor) gets activated directly via Notch signaling. TCF1 is a key regulator of T cell development; loss-of-function studies demonstrated profound defects with respect to cell survival and differentiation in both early and late stages of intrathymic development of T cells (Del Real and Rothenberg, 2013; Weber et al., 2011). TCF1 generally acts through Wnt-signaling, but it functions independently of Wnt pathway (Del Real and Rothenberg, 2013; Tiemessen et al., 2012) in determining T cell fate. TCF1 acts as a 'pioneer' factor and determines the chromatin landscape during early T cell development (Emmanuel et al., 2018; Johnson et al., 2018).

1.1.5.2 GATA3

A member of GATA family of zinc-finger transcription factor, GATA3 expression increases dramatically in response to Notch signaling at the ETP stage, post which it cohorts with the Notch signaling to block the B cell fate through the DN2a stage (Garcia-Ojeda et al., 2013; Scripture-Adams et al., 2014). In TCR αβ DPs, GATA3 promotes CD4 fate over CD8 SPs. These multiple roles of GATA3 are likely explained by its capacity to bind distinct genomic regions in pro-T vs the later stages of T cell development (Nakayama et al., 2017; Zhang et

al., 2012). In the periphery, GATA3 drives differentiation of mature CD4 T cells toward T helper 2 (Th2) subtype over T helper 1 (Th1) fate.

1.1.5.3 BCL11B

B cell lymphoma 11B (BCL11B) is a C2H2-type zinc finger protein. It is highly T cell-lineage specific transcription factor. Its expression initiates in the late DN2a stage, not at the ETP stage where TCF1 and GATA3 expression is induced. BCL11B is essential for multiple transitory aspects as the stages of the development progresses. Particularly, BCL11B is indispensable for thymocyte survival through β-selection, as shown in Bcl11b knockout mice (Kominami, 2012; Liu et al., 2010). Using Bcl11b conditional knockout specific to haematopoietic lineage, it was demonstrated that BCL11b deficiency leads to a developmental arrest at DN2-DN3 like stage (Ikawa et al., 2001; Li et al., 2010a; Longabaugh et al., 2017). Without BCL11B phase 1 pro-T cells (mainly DN2a) lose their T cell commitment program even in the presence of Notch signalling and initiate a NK or dendritic like program (Ikawa et al., 2010; Li et al., 2010b). In the chromatin context, BCL11B functions via direct binding to regulatory regions, and recruitment of chromatin-modifying factors (Ikawa et al., 2001) and is important for both activation and repression of its bound targets during T cell lineage choice (Ikawa et al., 2001; Longabaugh et al., 2017).

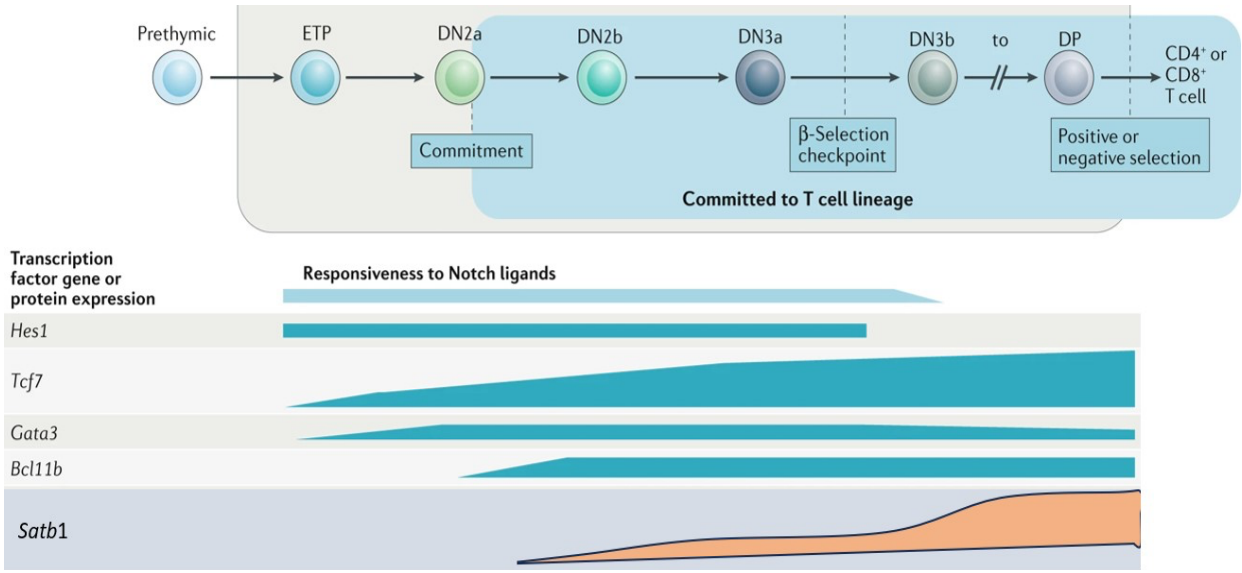


Figure 1.1.6: Major transcription factors involved in T cell development. Temporally controlled expression of various transcription factors is important for thymocyte progression through different stages. As early as ETP stage, Expression of TCF1 and GATA3 starts, followed by BCL11b as shown. There is a gradual increase in the expression of TCF1 as the

stages progress, indicating its importance in regulating distinct checkpoint. SATB1 exhibits basal expression in early stages upto DN3b, following which its expression is rapidly induced as a 'burst' until the DP stage. After the CD4/CD8 fate choice, the expression of SATB1 is modulated in accordance with the TCR signal received by the newly formed CD4/CD8 SP thymocyte. Adapted from Hosokawa and Rothenberg, 2021.

1.1.5.4 E proteins and ID proteins

E-proteins belong to basic helix loop helix (bHLH) class of transcription factors (Murre, 2019). E proteins' express from prethymic precursors stages (Dias et al., 2008; Miyazaki et al., 2017; Yang et al., 2008) and are implicated in both phase 1 (T lineage commitment) and later phases of development (β selection and DP fate choice) (Miyazaki et al., 2017), most strikingly during phase 2 where it induces the expression of multiple T lineage genes by modulating the chromatin landscape around the signature genes such as Rag1, Cd3e, CD3d, Cd3g, Notch1 and Ptcra (Miyazaki et al., 2017). E proteins function in either homo- or hetero-dimeric forms with other different E-proteins or other bHLH proteins to bind their consensus E-box containing DNA (Hosokawa and Rothenberg, 2021). E- protein E2A homodimer and E2A:HEB (encoded by Tcf3 and Tcf12, respectively) have been shown to have crucial roles early in T cell development (Bain et al., 1997; Wojciechowski et al., 2007). Tcf3 knockout results in a developmental arrest at the lineage commitment checkpoint (Ikawa et al., 2006). E2A is essential both in promoting NOTCH1 expression in prethymic precursors and cooperatively regulating Notch signaling targets to promote a T cell fate (Ikawa et al., 2006; Pereira de Sousa et al., 2012; Yashiro-Ohtani et al., 2009).

Since E proteins' expression (especially E2A) is not modulated drastically, given their requirement at different stages of T cell development, an inhibitory mechanism is timely placed to restrict their functionality. Antagonistic to E protein function are class IV HLH proteins called inhibitory to DNA binding (ID), which are devoid of essential DNA binding domain. ID proteins form stable E protein:ID heterodimers which are unable to bind E protein cognate targets. Notably, ID2 ectopic expression has been shown to drive pro-T cells toward an NK lineage fate (Ikawa et al., 2001), plausibly via inhibiting the DNA binding activity of E2A and HEB heterodimers (Miyazaki et al., 2017; Zook et al., 2018). E2A and HEB also play an important role in mediating TCR rearrangement (Braunstein and Anderson, 2012; Jones and Zhuang, 2011; Miyazaki et al., 2011).

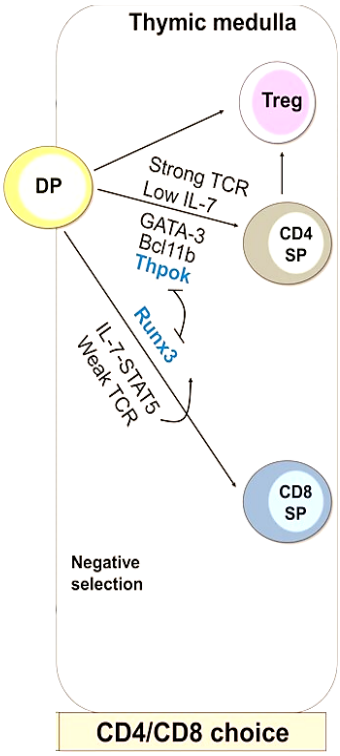


Figure 1.1.7: Lineage fate choice of CD4 versus CD8 by DP thymocytes. DP thymocytes receive different strengths of the TCR signal. Strong TCR signal promotes CD4 differentiation by the action of master-regulator GATA3 and TF BCL11B. These thymocytes uniquely express ThPOK, which represses the expression of Runx3, a CD8 lineage specifying protein. On the other hand, CD8 SP differentiation occurs when DP thymocytes receive a relatively weaker TCR signal or are poorly responsive to existing TCR signal (Kinetic model), Runx3 expression increases blocking ThPOK induction. Furthermore, IL7-STAT5 axis promotes CD8 specific transcriptional program. Reproduced from Sun et al., 2023.

1.1.5.5 Lineage fate determination by ThPOK and Runx3

ThPOK (encoded by Zbtb7b) is a transcription factor that functions as the master regulator in $CD4^+$ lineage choice (He et al., 2005). ThPOK expression is specifically repressed during $CD8^+$ lineage commitment (He et al., 2005; Sun et al., 2005). Loss-of-function or missense mutations cause complete lack of differentiated $CD4^+$ SP thymocytes (Dave et al., 1998; Egawa and Littman, 2008; He et al., 2005; Setoguchi et al., 2008; Wang et al., 2008), whereas overexpression of ThPOK biases DP thymocytes toward a $CD4^+$ cell fate (He et al., 2005). ThPOK is positively regulated by GATA3 and a TCR induced transcription factor SATB1 (Hernandez-Hoyos et al., 2003; Kakugawa et al., 2017; Pai et al., 2003). ThPOK additionally

antagonizes Runx3 and some CD8 specific genes to suppress commitment toward $CD8^+$ lineage (Egawa and Littman, 2008; Gao et al., 2022; Luckey et al., 2014). Contrastingly, Runx3 antagonizes $CD4^+$ lineage by directly repressing Cd4 and ThPOK gene expression (Taniuchi et al., 2002) and promotes $CD8^+$ lineage choice (Figure 1.1.7). Runx3 is upregulated in the DP thymocytes which receive low TCR strength/duration and is dependent on the Il7-STAT5 axis (Park et al., 2010).

1.1.5.6 Structural protein CTCF and the cohesin complex proteins in T cell development

CCCTC-binding factor (CTCF) is an evolutionarily conserved zinc finger protein. It could function as either transcription factor or repressor (Klenova et al., 1993). Multiple studies have demonstrated that the primary role of CTCF is anchoring the chromatin domain boundaries (Figure 1.1.8A) (Chen et al., 2008; Dunn et al., 2003) which defines the topologically associated domains (TADs), as discussed below in detail. In thymocytes, CTCF, although does not exhibit stage specific expression, plays an important role in maintaining cell cycle progression (Heath et al., 2008). Conditional knockout of Ctcf in T cell lineage suggests its role in transitioning from DN3 to DN4 as well as ISP to DP stages.

Cohesin is a multi-subunit complex which has been implicated in diverse biological processes including cell-cycle (Lengronne et al., 2006), DNA replication and DNA repair (Heidinger-Pauli et al., 2008; Strom et al., 2007). The cohesin complex is a ring-like structure comprised of Structural and Maintenance of chromosome 1 and 3(SMC1 and SMC3), RAD21 which is an

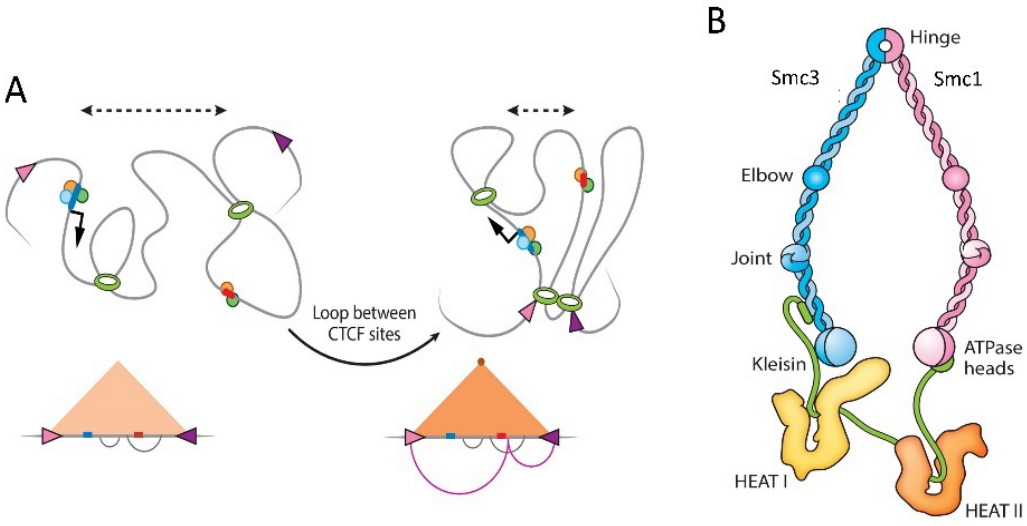


Figure 1.1.8: Genomic looping by Ctcf and the cohesin complex. A. Ctcf helps in creating a structural framework which ensures specificity of chromatin interactions over large distances leading to gene regulation. **B.** DNA loops are extruded by SMC complexes. The overall design of SMC complexes is shown. A pair of antiparallel coiled-coil subunits known as SMCs make

up the core of an SMC complex. These SMCs form a ring-like structure when they dimerize at one end (the hinge domain), and cooperatively interact at the opposite end (the ATPase heads). Reproduced from Kim et al., 2023.

α-kleisin subunit protein (Nasmyth and Haering, 2005). The fourth subunit of cohesion is either STAG1 or STAG2, which interacts with both SMC and RAD21 (Figure 1.1.8B) (Nasmyth and Haering, 2005) and facilitates appropriate interaction of cohesin with DNA (Cuadrado and Losada, 2020). It was shown that cohesin plays an important role in V(D)J recombination in αβ DP thymocytes (Seitan et al., 2013; Seitan et al., 2011; Seitan et al., 2012). Conditional knockout of Rad21 demonstrated that looping of different variable (Vα) regions onto the Jα chains is affected (Seitan et al., 2011), suggesting the role dramatic changes in chromatin contacts via cohesin could modulate T cell development.

1.1.5.7 SATB1

As the DP cells possess fully rearranged αβ T cell receptor, TCR signal potentiates expression of many lineage-defining TFs. Satb1 (Special AT-rich binding protein 1) is an early-TCR signal responsive TF (Galande et al., 2007; Gottimukkala et al., 2016; Pavan Kumar et al., 2006) and T cell enriched chromatin organizer. Satb1 is shown to be indispensable for appropriate thymocyte differentiation as Satb1-null thymocytes arrest at DP stage with few differentiated CD4 cells (Alvarez et al., 2000; Kondo et al., 2016). Studies have also shown that Satb1 is important for correct embryogenesis (Goolam and Zernicka-Goetz, 2017) and haematopoiesis (Satoh et al., 2013; Will et al., 2013).

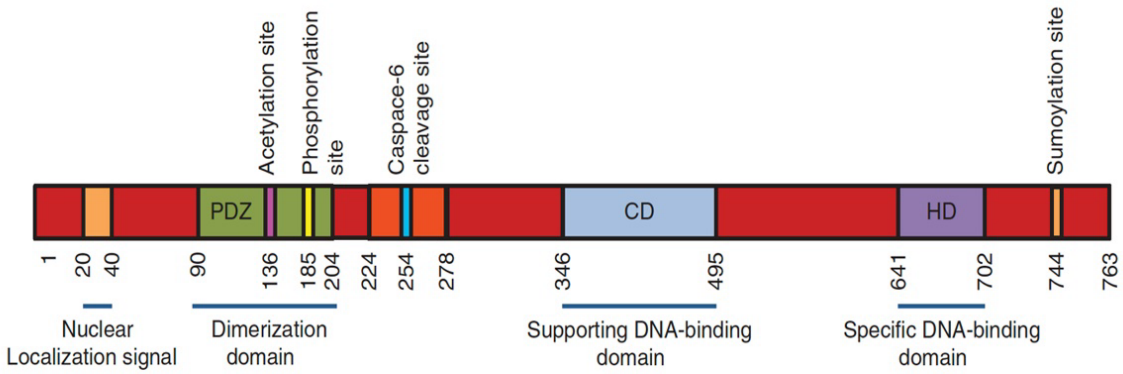


Figure 1.1.9: Schematic overview of Satb1 functional domains. Crucial amino acid residues and functional domains are displayed in this schematic depiction of the SATB1 protein. Dimerization requires the N-terminal ULD/PDZ domain, which has also been shown

to mediate interaction with different chromatin modifiers and transcription factors. The CUT domain is necessary for DNA binding of SATB1. Whereas, DNA-binding specificity is conferred by the homeodomain (HD). Reproduced from Burute et al., 2012.

Satb1 is originally characterized as a protein associated with the matrix attachment regions (MARs) (Dickinson et al., 1992). MARs are specific DNA sequences which possess high affinity to the nuclear matrix forming the base of the chromosomal folds and helping in condensation of chromatin. Structurally, Satb1 is composed of a PDZ or ubiquitin like domain (ULD) at the N-terminus, a CUT domain and a homeodomain (HD) at the C-terminus. CUT and HD of Satb1 in conjunction constitute the MAR binding region (Figure 1.1.9). PDZ/ULD domain of Satb1 is responsible majorly in mediating physical interactions with itself and other proteins. Satb1 functions by anchoring at the base of chromosomal loops, establishing a chromatin 'loopscape' (Figure 1.1.10) (Galande et al., 2007).

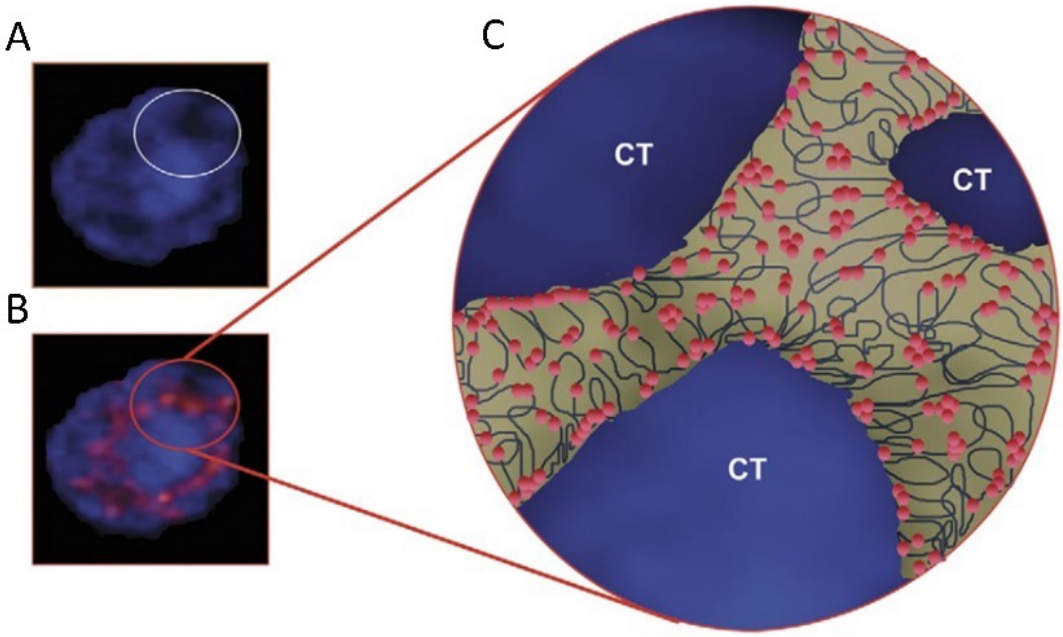


Figure 1.1.10: SATB1 mediated global chromatin looping. **A.** DAPI-stained mouse thymocyte display chromatin territories (CTs) which may represent heterochromatin **B.** SATB1 typically forms a "cage-like" network architecture around the CTs in thymocytes. **C.** Schematic illustration of an expanded thymocyte region showing SATB1 occupancy at different sites inside the chromatin. Reproduced from Galande et al., 2007.

Satb1 also plays critical role in establishing thymic regulatory T cell differentiation program by regulating super-enhancer (SE) of master regulator Foxp3 (Kitagawa et al., 2017). Due to its highest expression in T cell lineages, many studies focussed its multiple roles in DP differentiation (Kakugawa et al., 2017), autoimmunity (Kitagawa et al., 2017; Kondo et al., 2016), and T cell migration (Naito et al., 2023; Yasuda et al., 2019). Satb1 can act as both activator and repressor dependent on its post-translational modifications (Pavan Kumar et al., 2006; Purbey et al., 2009). Satb1 also recruits other chromatin modifiers such as PCAF/p300, HDACs (Cai et al., 2006; Kumar et al., 2005) and Swi-Snf complex (Yasui et al., 2002), as well as other TFs like β-Catenin (Mir et al., 2016; Notani et al., 2010) and Ctbp1 (Purbey et al., 2009) onto its genomic targets. Satb1 is also involved in T helper (Th2) CD4 differentiation (Ahlfors et al., 2010; Cai et al., 2006) as well as partially regulating lineage-specifying genes like Cd4, Cd8, Runx3, Zbtb7b (Kakugawa et al., 2017). Satb1 is also known to loop genomic regions at MHC Class-I locus (Kumar et al., 2007) and mediating anti-silencer to promoter interactions at *Rag1* locus (Hao et al., 2015) (Figure 1.1.11).

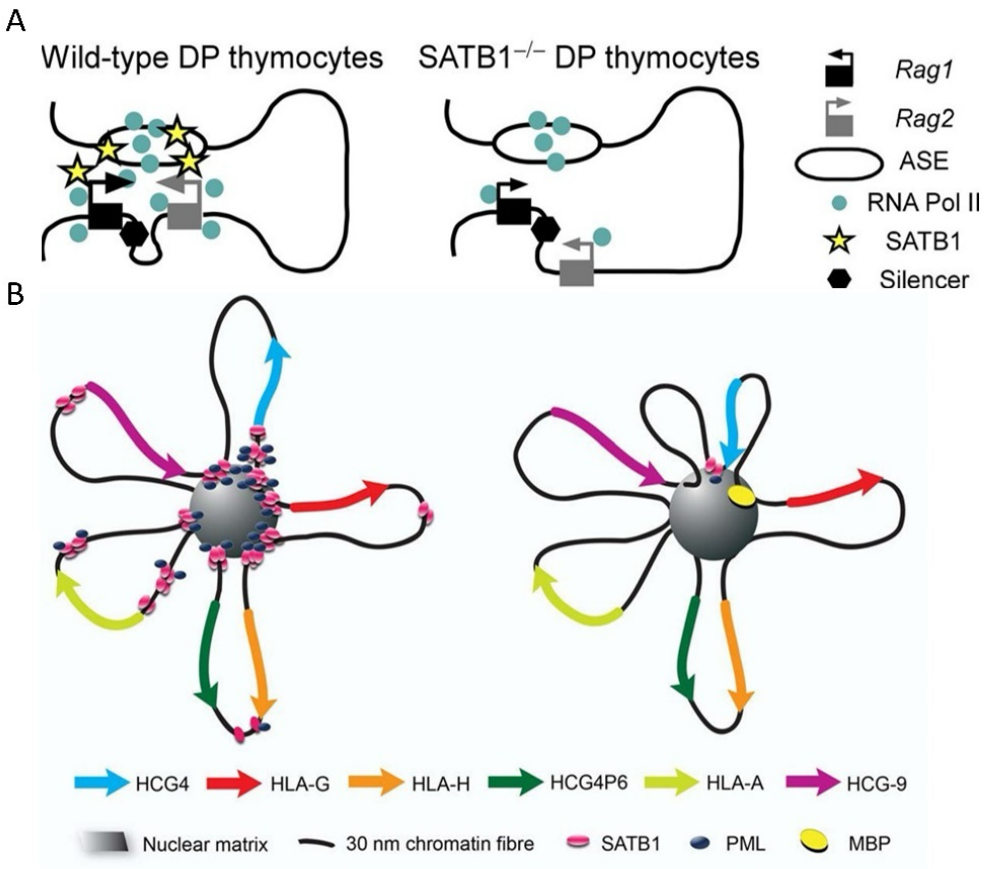


Figure 1.1.11: Satb1 mediated chromatin looping at Rag1/2 and MHC-I loci. A. In DN thymocytes, the RNA polII-loaded ASE has little interaction with the promoters of Rag1 and Rag2, which are translated at low levels. In DP cells, as SATB1 is levels are elevated it binds

to the anti-silencer element (ASE) and promoter of Rag1 transcription via Pol II. **B.** Diagram illustrating the MHC-I chromatin loop configuration in IFNg-treated and untreated cells depicting that the distribution of SATB1 and PML is important for specific loop structure of the MHC-I locus. Reproduced from Hao et al., 2015 (A) and Galande et al., 2007 (B).

1.1.6 Chromatin looping and gene regulation

The higher-order 3D structures called topologically associated domains (TADs) through globular interactions sequester regulatory elements (enhancers, promoters and non-coding RNAs), insulating them spatially from the neighbouring TADs or loop domains (Figure 1.1.12) (Dixon et al., 2015; Dowen et al., 2014; Nora et al., 2012). Recent accumulating reports support the concurrence of these self-assembled chromatin structures of the genome (TADs, sub-TADs) with regulation of gene expression and cellular identity (Stadhouders et al., 2019; Stadhouders et al., 2018; Zheng and Xie, 2019). Factors such as Ctcf and the Cohesin complex are shown to be enriched at the TAD boundaries (Figure 1.1.13) (Dixon et al., 2012; Fudenberg et al., 2016; Sexton et al., 2012), and have been implicated in formation of TADs via loop-extrusion (Alipour and Marko, 2012; Fudenberg et al., 2016; Sanborn et al., 2015).

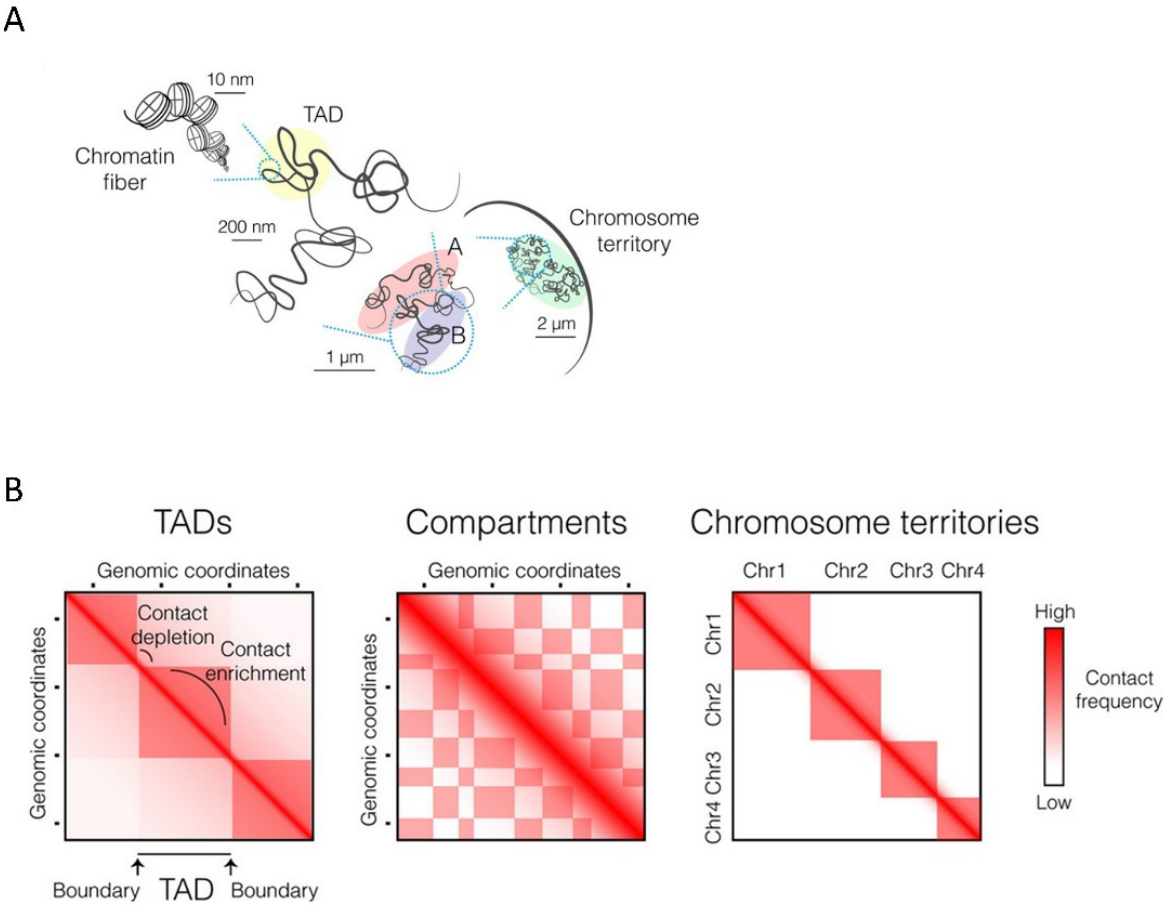


Figure 1.1.12: Organization of eukaryotic genome into chromatin domains. A. Diagram illustrating how chromosomes fold within the nucleus. The DNA-histone interaction is where the smallest layer of genome folding. The chromatin is packaged at various nucleosome densities. At the sub-megabase (<Mb) magnitude, the chromatin folds into topologically associated domains (TADs). There is a general chromosomal segregation of chromatin into active “A” and silenced “B” compartments of contacts. **B.** Hi-C maps schematically represented at various chromosomal sizes, indicating the many levels of higher-order chromosomal folding. TADs are shown as interaction-rich areas and are defined by TAD borders. Reproduced from Szabo et al., 2019.

Although of critical role, removal of Ctcf does not affect a substantial proportion of domain boundaries (Nora et al., 2017; Wutz et al., 2017), and affects minimally on global gene expression (Rao et al., 2017). Whereas knockdown of cohesin subunits showed ablation of

most of the loop domains, but strengthened genomic partitioning into compartment domains (Rao et al., 2017; Schwarzer et al., 2017).

More recent evidences show an instructive role of lineage-specific TFs in mediating the 3D interactions (Di Giammartino et al., 2019; Johanson et al., 2018; Weintraub et al., 2017), although the underlying mechanisms are still a subject of active investigation.

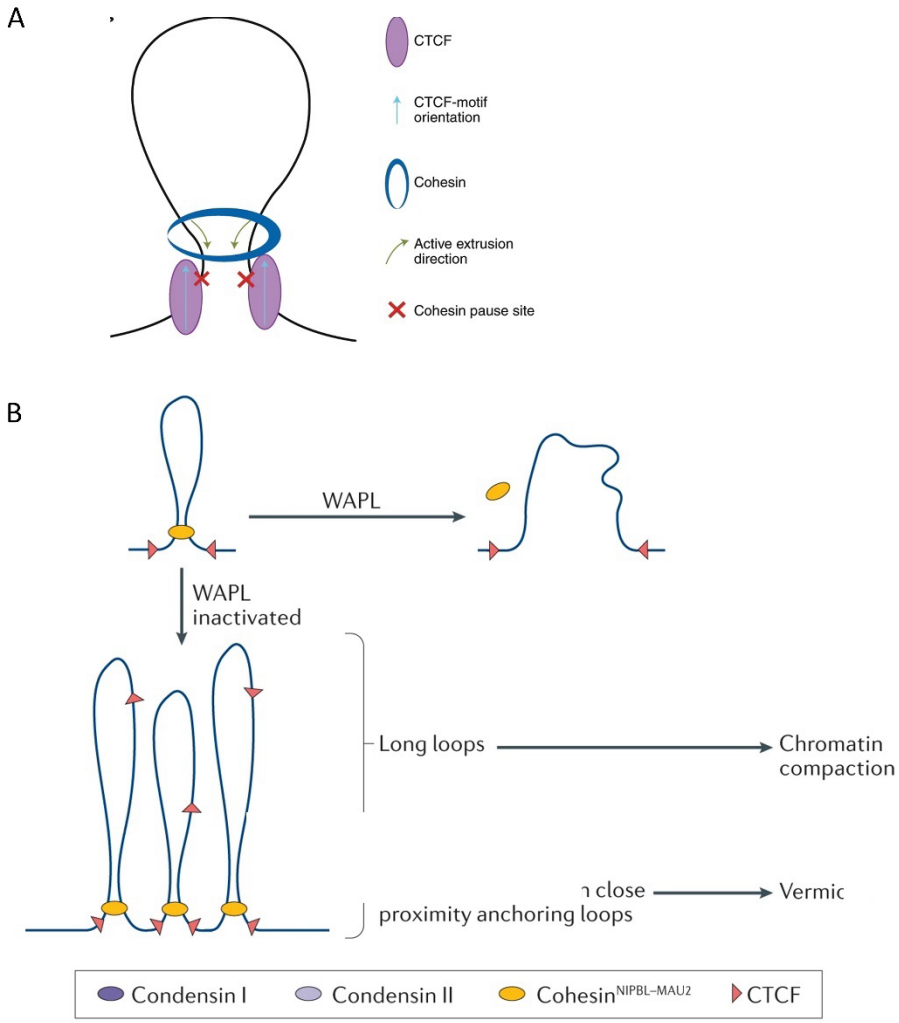


Figure 1.1.13: Cohesin mediated loop-extrusion as the molecular mechanism for genome folding. A. Topologically associated (TAD) loop domains are structurally based on contacts that are formed by the ATP-dependent extrusion of DNA via cohesin. **B.** Chromatin loops generated by cohesin are broken by WAPL-mediated disengagement of cohesin from the chromatin. When WAPL is inactivated, cohesin is prevented from releasing, which leads to the creation of larger loops known as 'vermicelli' -as indicated by higher compaction of the chromatin. Reproduced from Beagan and Phillips-Cremins, 2020 (A) and Davidson and Peters, 2021 (B).

About 10-40% TAD boundaries are variant depending on the cell-type, endowing plasticity to the 3D architecture and lineage-specific transcription (Bonev et al., 2017; Pekowska et al., 2018). T cell development is one of the ideal models to capture the extent of such variation (Singh et al., 2014).

Very recently the role of Satb1 in maintaining DP identity via regulating DP specific SE (Zelenka et al., 2022) and genomic structural reorganization at specific loci (Zelenka et al., 2022) has come to light. Despite the multifaceted roles of Satb1 described its role during DP fate choice is still only partially understood.

1.2 Materials and Methods

1.2.1 Mice

Six- to ten-week-old WT Satb1 fl/fl, CD4cre Satb1 fl/fl and VaviCreERT2 Satb1 fl/fl mice were used to prepare the single cell suspension of thymocytes for the flow cytometry analysis as well as for sorting the subpopulations of thymocytes which are in different stages of T-cell development. All the strains were bred and maintained in sterile environment and the dissections were carried out in accordance with the guidelines of NFGFHD, IISER Pune.

1.2.2 Flow cytometry

The suspension of thymocytes and or splenocytes were prepared using thymii from 6-week-old mice from various genotypes indicated and were used for surface staining. Before staining, Fc -blocking was done with anti-CD16.32 antibody (BD Biosciences). Then, the thymocytes were subjected to surface staining: BV 421 anti-mouse CD4 (Clone GK1.5, Biolegend); APC anti-mouse CD3 (Clone GK1.5, eBioscience); PE-Cy7 anti-mouse CD69 (Clone GK1.3, BD Biosciences); PE anti-mouse CD8a (Clone 53-6.7, BD Biosciences); FITC TCRβ (Clone M1/69, eBioscience). For Splenic cells, APC anti-mouse CD62L (Clone 19.3, eBioscience) and PE-Cy7 CD44 (Clone B21.1, Biolegend) were used in combination with CD4 and CD8a staining. The analysis was performed in BD Celesta or BD Aria III Fusion SORP. The thymocyte sub populations such as CD4-CD8- double negative (DN), $CD4^+CD8^+$ double positive (DP), $CD4^+$ SP thymocytes (total CD4SP), and $CD8^+$ SP thymocytes (total CD8SP) were FACS sorted using FACS Aria III SORP (BD biosciences). The splenic CD4 subpopulations; activated and naïve were sorted based on CD44 and CD62L markers using FACS Aria III SORP (BD Biosciences).

1.2.3 Datasets and bioinformatics analysis

ChIP-seq studies of the genome-wide occupancy of H3K4me3, H3K27me3, and H3K4me1 in mouse DP thymocytes were conducted using the GSE20898 (Wei et al., 2011) publicly available data set. Datasets for Satb1, Smc1a and Ctcf ChIP seq were obtained from accession numbers GSM1617950, GSE41743 and GSE61428. For HiChIP of Satb1 and HiC in Satb1 WT and cKO conditions, we used the datasets from GSE173446. Bowtie2 was used to align the raw readings, and MACS2 was used to name peaks. The IGV genome browser was used to see the peaks that MACS created. R software ChIPseeker was used to plot genome-wide occupancy profiles and biological processes. The GSE48138 dataset was used to conduct RNA-seq analysis on DP and CD4SP thymocytes. Datasets from GSE182995 was used for ATAC-seq and HiC analysis as well as for H3K4me3 and H3K27me3 modifications; both in WT and Satb1 cKO conditions. GSE77695 data set was used for the analysis of ATAC-

seq performed in HSCs, MPPs and CLPs. GSE100738 data set was used for ATAC-seq analysis performed in DN1, DN2a, DN2b, DN3, DN4, DP, CD4SP and CD8SP. IGV v2.1 was used to plot the occupancies of ChIP-seq and ATAC-seq processed data onto selected genomic regions. HISAT2 and Stringtie were used to align RNA-seq reads. Additional differential gene expression analysis was conducted using R tool Deseq2.

1.2.4 Cloning

Satb1 and its specific domains PDZ/ULD and MD+HD were cloned into pTRIEX-3-neo to create C-terminal HIS tag as a fusion ORF with each domain. Full length Smc1a and its identified domains N-term, Coil-I, Hinge, Coil-II and C-term were cloned in pGEX4T1 upstream of the glutathione S transferase (GST). For CRISPR, gRNA sequences for Smc1a and Satb1 were obtained by using 'benchling' CRISPR tool. The top hits were added with Bsmb1 restriction sites and used for oligo synthesis. The complimentary oligos for each gRNA were annealed together at 95°C for 5 min and cooled slowly till 25°C. The annealed oligo was ligated into linearized (with BsmbI) LentiCRISPR v2.

1.2.5 Cell culture, transfection and reporter assay

Growth media supplemented with 10% fetal bovine serum (Gibco) was used to cultivate HEK 293T cells, which were then kept in a humidified incubator at 37°C and 5% CO2. Transfection of packaging vectors pMD2 and pPAX2, along with lentiCRISPR v2 cloned with individual gRNA was performed using Lipofectamine 3000 (Invitrogen). After 48 and 72 hrs the virus containing media was collected and filtered using 0.45µm syringe filter and stored in -80℃ for future use. For the reporter assay, HEK293T cells were transfected with IL2 promoter cloned in pGL3 basic. Alongside, either empty vector, SATB1, SMC1A or SMC1A+SATB1 were cotransfected. An equal amount of Renilla expressing pTRK vector was also transfected in each well. Cells were harvested after 48 hrs and subjected to passive cell lysis. Dual luciferase assay was performed using the manufacturer's protocol (Promega).

1.2.6 Lentiviral knockout of Satb1 and Smc1a in DP thymocytes

To isolate the thymus, C57BL/6 mice from 6 weeks old were used. A single cell suspension was made by crushing the thymus through a 70-micron filter. The following fluorochrome-tagged antibodies were then used to surface stain thymocytes: PE anti-mouse CD8a (Clone 53-6.7, BD Biosciences) and FITC anti-mouse CD4 (Clone GK1.5, BD Biosciences). Thymocytes that were double positive (DP) and $CD4^{+}CD8^{+}$ were sorted using the FACS Aria III SORP (BD Biosciences). To the sorted DP, lentivirus particles containing empty vector, Satb1 gRNA, Smc1a gRNA, or Satb1 and Smc1a gRNA were introduced. For 36 hours, thymocytes were grown in RPMI-1640 medium supplemented with 10% FBS, penicillin-

streptomycin, and IL7 at 5ng/mL, SCF at 5ng/ML, and IL2 at 2ng/ML. Following incubation, cells were taken out and subjected to flow cytometry analysis, western blotting or RNA isolation.

1.2.7 Chromatin Immunoprecipitation (ChIP) sequencing

DP thymocytes from WT and Satb1 cKO mice were used for chromatin immunoprecipitation (ChIP). For ten minutes, 50×10^6 sorted DP thymocytes were cross-linked at room temperature using 1% formaldehyde which was quenched using 125 mM glycine. Buffer 1 (25 mM Tris-HCl pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.1% NP40, 1 mM DTT, 0.5 mM PMSF and 1X protease inhibitor cocktail (PIC) (Roche) was used to isolate the nuclei from the cross-linked cells. The nuclei were resuspended in lysis buffer (50 mM Tris-HCl pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X–100, 0.1% Sodium deoxycholate, 0.1% SDS, 0.5 mM PMSF, and 1X PIC). Sonication of the chromatin was performed to get an average fragment distribution of 200–600 base pairs. The chromatin was then subjected to sonication using a Covaris M220 ultrasound sonication system for 12 min, peak power 75 and burst cycle of 200. Protein A/G magnetic bead slurry (Thermo Scientific) was used to preclear the sonicated chromatin for 1 hour. The beads were discarded and the precleared chromatin was subjected to antibody binding using anti-mouse SMC1A (Abcam) antibody overnight at 4°C. The antibody-bound protein (and associated chromatin) was precipitated by adding Protein A/G beads (saturated using 10 mg/mL tRNA and 1% BSA and at 4°C for 4 hr). The lysate was discarded and the beads were washed thoroughly using the lysis buffer. Elution of complexed from A/G beads was done using the elution buffer (1% SDS, 0.1 M $NaHCO_3$). The eluted chromatin was de-crosslinked and treated with the proteinase K and RNase A. Further the DNA was purified and subjected to either library preparation using manufacturers protocol or quantitative RT-PCR (qRT-PCR) analysis using the Sybr green qRT-PCR master mix (Takara). The following primers were used for qRT-PCR:

Gene	Sequence
Cd3e_promoter F	CTGAACCCCACACAGGAAGT
Cd3e_promoter R	TTTGCACAGGGCACAAGTAG
Cd8a_promoter F	TATGGCTTCATCCCACAACA
Cd8a_promoter_R	GACTGGCACGACAGAACTGA
Cd4_promoter_F	CTAGGCCCAAGAAGGTTTCC
Cd4_promoter_R	GGGAGAGAAAGTGCCTACCC

For ChIP-seq libraries were prepared using NEB-Next UltraII library prep kit (NEB) with an equivalent quantity of DNA (~5 ng) as input. Quantitative PCR (qPCR) analysis was used to estimate the number of cycles required for the amplification of adapter-ligated libraries. Using the HiPrep PCR clean up system (MagBio Genomics, USA), final libraries were purified. Prior to pooling libraries at an equimolar ratio, the library concentration was ascertained using Qubit, and the average fragment size was measured using the dsDNA HS test on the Bioanalyzer 2100 (Agilent). The HiseqX platform was utilized to acquire sequencing reads (150 bp PE).

1.2.8 RNA-seq

Under different conditions, RNA from cultured DP cells were extracted using a TRIzol for Bulk RNA-Seq. The Agilent 2100 Bioanalyzer and NanoDrop2000 were used to quantify and qualify the total RNA. Following this, 1 µg of total RNA was utilized to prepare the library. The Poly(A) mRNA Magnetic Isolation Module was used to carry out the poly(A) mRNA isolation. First strand synthesis mix (NEB) was used to create first-strand cDNA, while Second Strand Synthesis Enzyme Mix (NEB) was used to create second-strand cDNA. Three steps were taken with the isolated double-stranded cDNA: adapter ligation, 3'-dA tailing, and end repair. Cycling was determined using qPCR. Using Illumina P5/P7 primers, six to ten PCR amplification cycles were performed to amplify the ligated DNA. Beads and a Qubit 4.0 Fluorometer (Thermo Scientific) were used to clean up the PCR products. Libraries were pooled and sequenced using Illumina HiSeqX platform in a 2x150 bp paired-end mode.

1.2.9 cDNA synthesis and Quantitative PCR analysis (qRT-PCR)

RNA from DP thymocytes sorted from WT and Satb1 KO mice was performed using RNeasy mini kit (Qiagen). After DNase I (Roche) treatment, the RNA was cDNA synthesis was done using High-capacity cDNA synthesis kit (Applied Biosystems). Quantitative PCR was carried out with qPCR master mix (Takara) using the below PCR conditions: 95°C-5 min, 95°C-30 sec, 60°C-30 sec, 72°C-45 sec -for 37 cycles, 72°C-3min. The following qPCR primers were used:

Gene	Sequence
Cd8a F	CTTCCAGAACTCCAGCTCCA
Cd8a R	ACCGAGTTGCTGATGACTGA
Cd4b F	TTTGCAGAGGAAAACGGGTG
Cd4 R	AGAGTCAGAGTCAGGTTGCC
Cd3e F	CCTGAAAGCTCGAGTGTGTG
Cd3e R	TGGGCTCATAGTCTGGGTTG

1.2.10 Assay for transposase-accessible chromatin (ATAC) -qPCR

To prepare for ATAC reaction, 50,000 cells were twice washed with PBS, lysed in 50 µL of ice-cold Lysis Buffer (50 mM HEPES, 0.1% NP40, 50 mM NaCl). The isolated nuclei were then pelleted for ten minutes at 500 X g at 4°C. 100 nM Tn5 Transposase was used to execute the tagmentation process in 1X Tagmentation Buffer (10 mM Tris pH 7.4, 5 mM $MgCl_2$, 10% DMF, 33% PBS, 0.1% TWEEN 20, 0.01% Digitonin) for 30 min at 37°C. DNA was purified using PCR Purification columns (MN biosciences). Before qPCR, DNA was eluted into a volume of 20 µl and kept at -20°C. After amplification using one set of i5 and i7 primers (NExtera), qPCRs were performed for *Cd3e* promoter using the ChIP primers.

1.2.11 Co-Immunoprecipitation Immunoblotting

Thymocyte single-cell suspension was made from thymi isolated from six-week-old mice. Thermo Scientific's BCA kit was used to assess the protein content after cells were lysed using NP40-based lysis buffer. Using 3 µg mouse or rabbit IgG and protein A/G Dyna beads (Thermo Scientific), 500 µg of protein was precleared for 1 hr. Using anti-SATB1 antibody, the cleared lysate was subjected to antibody-binding for four hours at 4°C. Following incubation, protein A/G magnetic beads (Thermo Scientific) were used to precipitate anti-SATB1-bound protein complexes. At least 3 washes were given to the beads with lysis buffer. 1% SDS in 1X PBS was used to elute proteins at 98°C for 15 min. Co-Immunoprecipitated proteins were identified by subjecting the eluted protein to SDS gel electrophoresis followed by western blotting of SMC1A.

1.2.12 Western Blotting

BCA method was used to quantify the proteins after the cells were lysed using RIPA lysis buffer (50 mM Tris-HCl at PH 7.4, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, 0.25% sodium deoxycholate, 1 mM PMSF, and 1X protease inhibitor cocktail (Roche). After the whole protein was separated on a 10% SDS polyacrylamide gel, it was transferred to a PVDF membrane (Millipore). Blocking of non-specific interactions was done with 5% skimmed milk. The following antibodies were used to probe the blots: anti-Actin (1:1000, Millipore), anti-SATB1 (1:1000, Santacruz Biotechnologies), and anti-Smc1a (1:1000) at 4°C overnight. Next day, blots were washed 5 times with 1X TBST (500 mM NaCl, 20 mM Tris pH 7.5, 1 mM EDTA, 1% Tween 20) and anti-primary secondary antibodies tagged with HRP (anti-mouse IgG-HRP or anti-rabbit-HRP) were used to probe the blots followed by extensive washing with 1x TBST. Signals were developed using the ECL luminescence detection reagent (BioRad) and captured using the ImageQuant LAS 4000 system (GE Healthcare Life Sciences) and.

1.2.13 Immunostaining

A single-cell suspension was obtained from thymi of WT and Satb1 cKO mice that were six weeks old. In order to fix thymocytes, 2% paraformaldehyde was used. Permeabilization of fixed thymocytes was performed with 0.1% Triton X-100. Following permeabilization, thymocytes were treated for three hours at room temperature with mouse anti-SATB1 (Abcam) and rabbit anti-SMC1A (Abcam). Thymocytes were stained intracellularly and then washed with 1X PBS containing 0.01% Tween 20. The secondary antibodies were stained for one hour at room temperature using the fluorochrome-tagged secondary antibodies, anti-mouse GFP 488 and anti-rabbit Alexafluor 594. Following the staining with a secondary antibody, DNA was labeled with DAPI (Sigma). Z-stack pictures were obtained and cells were seen using a Sp8 microscope (Leica Biosystems).

1.2.14 HiC and HiChIP analysis

After adapter cutting and quality filtering, reads were acquired using fastp. Hi-C mapping was carried out using HiC-Pro. The mm10 mouse reference genome was mapped to the paired-end reads. HiC-Pro was used to filter valid reads after reference mapping. At various resolutions such as 10-kb, 100-kb, and 1 Mb, raw contact matrices were created. HiC-Pro (v2.11.1) was employed to the raw contact matrices from Hi-C data. The .hic files were then generated using HiC-Pro Using Juicebox, Hi-C contact matrices were visualized. Similar processing was used for HiChIP data. HiChIP data was plotted using HiGlass software.

1.3 Results

1.3.1 Satb1 co-occupies genomic regions with Ctcf and Cohesin in DP thymocytes

It has been shown that Ctcf binding sites, when specifically in a convergent orientation generally marks a TAD boundary (Wendt et al., 2008). The current leading mechanism that governs domain formation is the loop extrusion model, according to which structural maintenance of chromosomes (SMC) containing Cohesin complex (A ring-like complex) when loaded on to the genome, it could track on the DNA and extrudes the passing DNA into loops (Nasmyth, 2001; Riggs, 1990). While a vast milieu of literature focused on Ctcf mediated boundary formation and TAD organization, perturbations to which have shown a little to moderate impact on cell-type specific gene regulation. Therefore, to identify a cell-type specific factor whose expression is stage-dependent and could mediate genomic looping, we chose Satb1, expression of which is dependent on the TCR signal. To understand the functional significance of Satb1 mediated genome organization during T cell development, we first employed publicly available ChIP-seq and Hi-ChIP data for Satb1 as well as HiC measurements in the DP stage of thymocyte development at which Satb1 expression peaks (Immgen. Database, Figure 1.3.1A). We overlayed HiChIP loops of Satb1 with the ChIP seq tracks of Satb1, Ctcf and histone modification H3K27ac. We observed that Satb1 mediated loops are highly correlated with Ctcf binding (Fig 1.3.1B). H3K27ac also shows a strong overlap with Satb1 binding as exemplified at the *Ets1* locus (Fig 1.3.1B). ChIP-seq analysis of Satb1 and Ctcf revealed >70% common binding sites among the two and a strong overlap in their average TSS binding (Fig 1.3.1C). Ctcf is known to regulate TAD formation via restricting Cohesin mediated loop-extrusion and is enriched on TAD boundaries in thymocytes (Beagan and Phillips-Cremins, 2020). We called DP specific TAD boundaries, to gain the interaction coordinates. Upon probing for Satb1 binding peaks within the boundaries, we find that Satb1 also occupies ~22% TAD boundaries (Fig 1.3.1D), and binds 3448 of 6093 boundary associated genes (Fig 1.3.1E), which potentially suggests its role in looping DNA in a stage-specific manner in tandem with the constitutive action by Ctcf. Furthermore, at specific immune genes' promoters and enhancers, Satb1 could bind exclusively of Ctcf, or commonly with Ctcf as exemplified by the Cd8 SE region (Fig 1.3.1F). These results suggest that Satb1 could bind chromatin in DP thymocytes in conjunction as well as exclusively of Ctcf, and may have a role in modulating chromatin architecture in DP thymocytes.

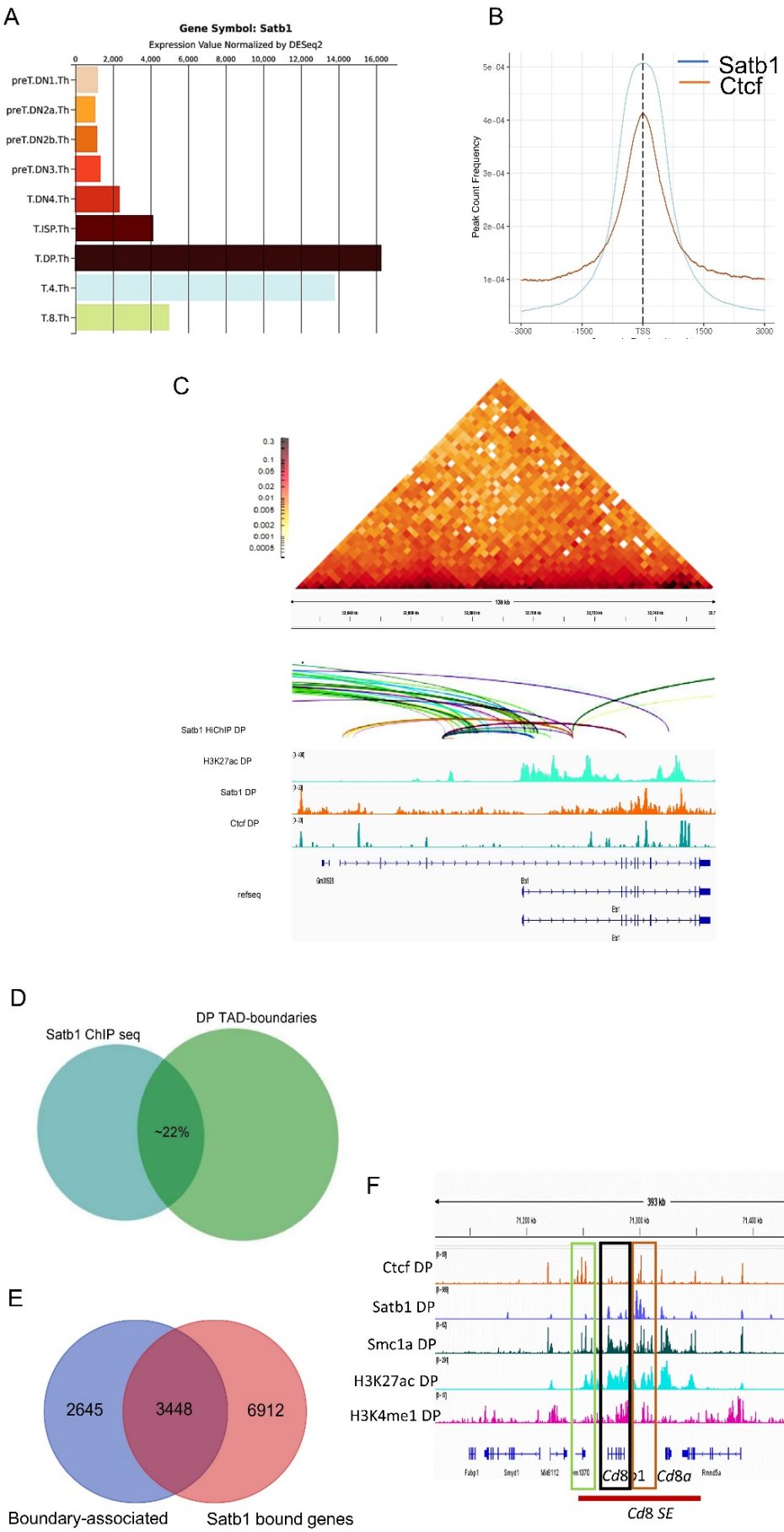

A
Gene Symbol: Satb1
Expression Value Normalized by DESeq2
0
2,000
4,000
6,000
8,000
10,000
12,000
14,000
16,000
preT.DN1.Th
preT.DN2a.Th
preT.DN2b.Th
preT.DN3.Th
T.DN4.Th
T.ISP.Th
T.DP.Th
T.4.Th
T.8.Th
B
Satb1
Ctcf
Peak Count Frequency
5e-04
4e-04
3e-04
2e-04
1e-04
-3000
-1500
TSS
1500
3000
C
0.3
0.1
0.05
0.02
0.01
0.005
0.002
0.001
0.0005
Satb1 HiChIP DP
H3K27ac DP
Satb1 DP
Ctcf DP
refseq
D
Satb1 ChIP seq
DP TAD-boundaries
~22%
E
2645
3448
6912
Boundary-associated genes
Satb1 bound genes
F
Ctcf DP
Satb1 DP
Smc1a DP
H3K27ac DP
H3K4me1 DP
Cd8b1
Cd8a
Cd8 SE

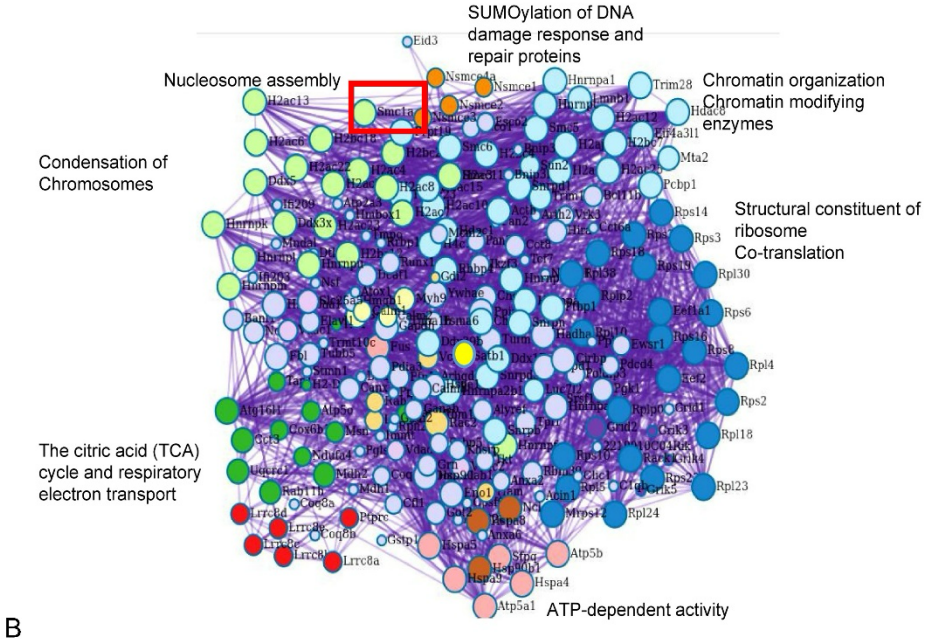
Figure 1.3.1: Satb1 co-occupies genomic regions with Ctcf and Cohesin in DP thymocytes. A. RNA-seq analysis of various populations of the developing T cells using Immgen database, showing the induction stages and transcript levels of Satb1. **B.** Publicly available ChIP seq data of Satb1 and Ctcf in DP thymocytes were analysed. Average chromatin occupancy profiles of Satb1 and Ctcf at gene promotors/TSS in DP thymocytes are shown. **C.** Satb1 HiChIP, WT HiC and ChIP seq data for H3K27ac, Satb1 and Ctcf were analysed and plotted together in an IGV track, depicting Satb1 mediated loops linking Ets1 promoter and enhancer. Satb1 binding also conforms highly with Ctcf and H3K27ac at various loci. **D.** HiC analysis of WT DP thymocytes was carried out and TAD boundaries were called. Separately Satb1 ChIP-seq analysis was done and unique peaks were called. The number of common regions between the two sets is shown, demonstrating Satb1 could demarcate specific TAD boundaries. **E.** Similarly nearest genes from both the sets obtained in **D** were called and common genes between the two gene sets are shown here. **F.** Common as well as exclusive genomic occupancy of Satb1 and Ctcf is exemplified at the Cd8 super-enhancer (SE) loci. ChIP-seq data for the indicated factors were analysed and genomic tracks were plotted together to observe enrichment at various immune specific genes.

1.3.2 Satb1 physically interacts with Cohesin in DP thymocytes

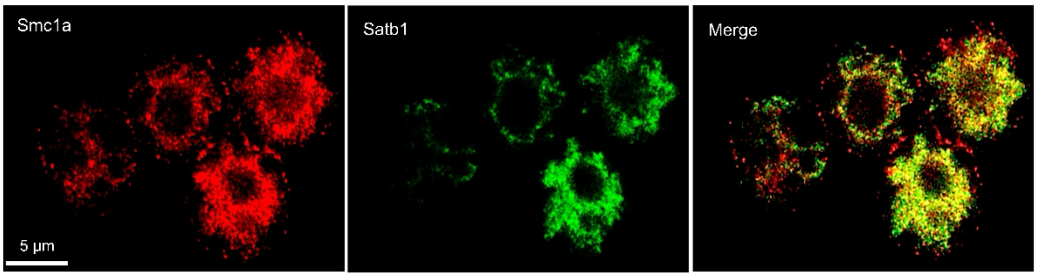
In previous studies from multiple groups, it was shown that Satb1, having specific affinity to base-unpairing AT-rich sequences (Dickinson et al., 1997; Dickinson et al., 1992), sequester different TFs (Kumar et al., 2007; Notani et al., 2010) and chromatin modifiers (Cai et al., 2006; Pavan Kumar et al., 2006; Purbey et al., 2009; Yasui et al., 2002) to its genomic targets. Both early and recent studies demonstrated that different factors might collaborate to mediate chromatin architecture. For instance, Mediator and Cohesin complexes have been shown to link genomic architecture to transcriptional activation in certain contexts (Kagey et al., 2010). Similarly, to investigate the involvement of other chromatin-associated factors with SAtb1, we performed immunoprecipitation mass-spectrometry (IP-MS) of Satb1 using DP thymocyte cell-lysate. With 90% confidence, we identified ~250 interacting proteins with significant enrichment. Using Cytoscape (Reimand et al., 2019), we performed PPI network analysis ($p<0.05$) and clustered them into biological processes (Fig. 1.3.2A). Most prominent clusters were Chromatin Organization and Chromatin Modifiers, Nucleosome assembly, and Co-translation. Of interest, we observed a high enrichment of Cohesin complex protein Smc1a in our dataset. Smc1a is an evolutionarily conserved core subunit of the Cohesin complex (Kim et al., 2023), which is involved in chromosomal segregation and 3D genome organization (Rao

et al., 2017; Seitan et al., 2012). To confirm the interaction, we performed super-resolution confocal microscopy for Satb1 and Smc1a immunostaining in DP thymocytes (Fig 1.3.2B). Satb1 has been shown to display an unusual ‘cage-like’ architecture in thymocytes (Notani et al., 2010; Yasui et al., 2002). Interestingly, we observed that Smc1a also forms this distinct immunostaining pattern (Fig 1.3.2B). Z-sectioning revealed that Satb1 and Smc1a co-localize throughout the depth of the nucleus (Fig. 1.3.2C) in distinct nuclear puncta. Furthermore, we performed co-Immunoprecipitation of Satb1 with Smc1a pull-down and observed a strong interaction between the two (Fig. 1.3.2D and E).

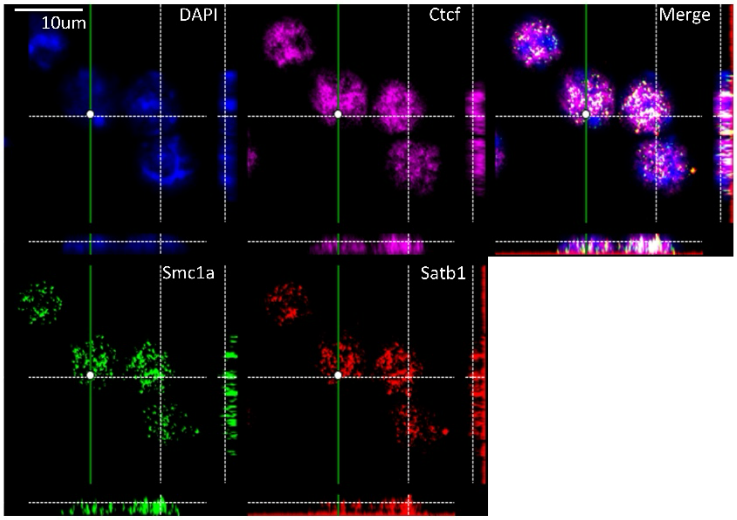
A


SUMOylation of DNA damage response and repair proteins
Nucleosome assembly
Chromatin organization
Chromatin modifying enzymes
Condensation of Chromosomes
Structural constituent of ribosome
Co-translation
The citric acid (TCA) cycle and respiratory electron transport
ATP-dependent activity

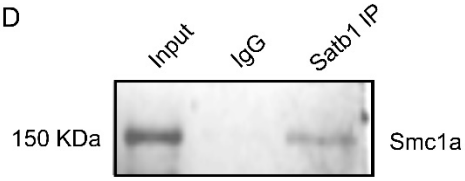
B


Smc1a
Satb1
Merge
5 μm

C


10um
DAPI
Ctcf
Merge
Smc1a
Satb1

D


Input
IgG
Satb1 IP
150 KDa
Smc1a

E

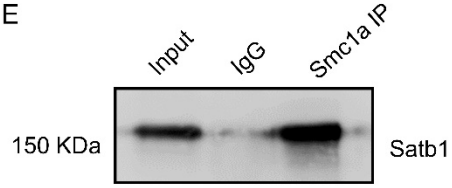

Input
IgG
Smc1a IP
150 KDa
Satb1

Figure 1.3.2: Satb1 physically interacts with Cohesin in DP thymocytes. A. Protein lysate from DP thymocytes was prepared. Immunoprecipitation (IP) of Satb1 was carried out and mass-spectrometry (MS) was done on the IP-ed proteins. Proteins enriched significantly (>90%) with at least 3 peptides were selected for the protein network analysis. The network displayed shows the most significantly enriched processes in which Satb1 interactor proteins are involved. Smc1a, highlighted in red, was detected as a Satb1 interacting protein. **B.** Co-localization of Satb1 and Smc1a was confirmed in DP thymocytes using super-resolution immunostaining. Smc1a was labelled with Alexa-fluor 594, shown in red. Satb1 was labelled with Alexa-fluor 488, shown in green. The merged image shows their interaction frequency, depicted as yellow points, which displays a unique 'cage-like' pattern in DP thymocytes. **C.** Immunostaining of Satb1 (Red), Smc1a (green), Ctcf (Magenta) and DAPI (blue) was carried out. z-sectioning revealed that Satb1 and Smc1a co-localize throughout the nuclear depth and form distinct 'foci-like' staining patterns, unlike Ctcf. **D** and **E.** Satb1 interaction was confirmed with Smc1a via co-immunoprecipitation (Co-IP). Satb1 IP was performed, followed by Smc1a western blotting (WB), and vice-versa.

1.3.3 Dissection of interaction domains of Satb1 and Smc1a

Satb1 consists of 2 major regions -a PDZ or a ULD N-terminal domain and a DNA binding arm (MD+HD) (Pavan Kumar et al., 2006). The DNA binding region is further classified as Cut1 and Cut2 domains, stabilized onto the DNA via a Homeodomain (HD) (Uniprot:Q01826). Majority of protein-protein interactions by Satb1 involved the PDZ/ULD domain. Whereas the Cut1 and Cut2 bind the DNA in a AT-rich context, and HD provides sequence specificity (Purbey et al., 2008). Satb1 can form homodimers and homotetramers under physiological conditions (Galande et al., 2001; Wang et al., 2012).

SMC1A has five distinct domains consisting an N-terminal, a C-terminal, two coiled-coil domains and a hinge domain connecting the two coil domains (Uniprot: Q14683). The N terminus contains a nucleoside triphosphate (NTP) binding region (Walker A box), whereas the C terminus is responsible for its DNA binding (containing a Walker B motif). The hinge region of Smc1a provides flexibility as well as interaction with other proteins, mainly the second core subunit SMC3 of the cohesin complex (Musio, 2020). RAD21 stabilizes the complex via binding the 'Head domain' (containing N and C terminus), this forming a ring-like structure.

To identify the regions of physical interaction for Satb1 and Smc1a, we created deletion constructs of Satb1; ULD/PDZ-like, Cut1+Cut2+HD with HIS tag (1.3.3A and Smc1a; N-term,

Coil-I, Hinge, Coil-II and C-term as GST fusion proteins (1.3.3B). Using purified proteins of full-length Satb1 (1.3.3C) and Smc1a, and their domains (1.3.3D), we performed in vitro pull-down assays. Upon incubating immobilized HIS-Satb1 with individual Smc1a domains, we observed that Satb1 interacts strongly with Coil-II and C-term domains of Smc1a whereas a weaker interaction is observed with the hinge region (1.3.3E). Whereas Smc1a interacts exclusively with ULD/PDZ-like domain (1.3.3F), which is the predominant protein-protein interaction domain for Satb1. Collectively, these results suggest that Satb1 physically interacts with the Cohesin subunit Smc1a, and their interaction may primarily be dependent on their N-terminal and C-terminal regions, respectively.

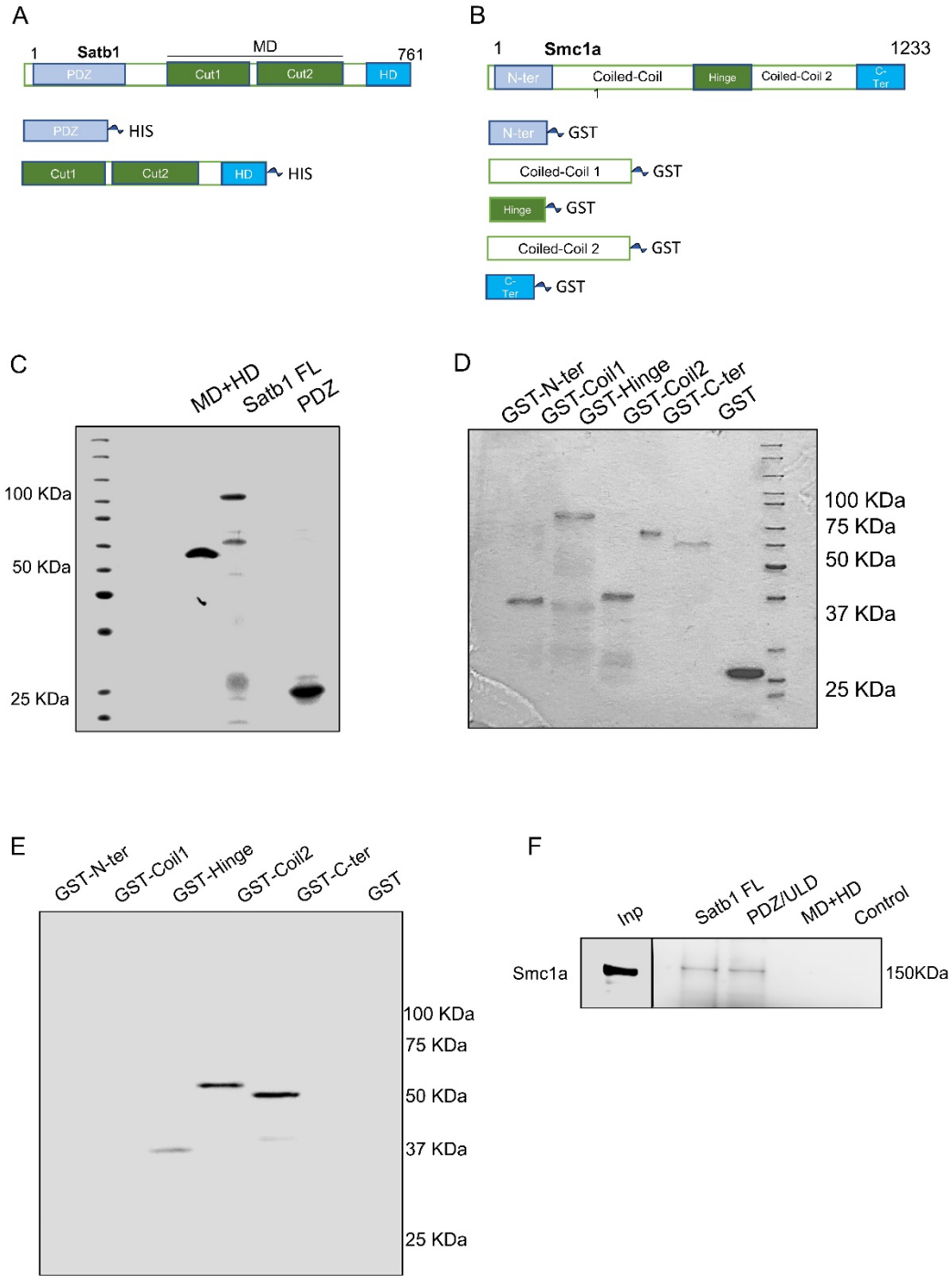
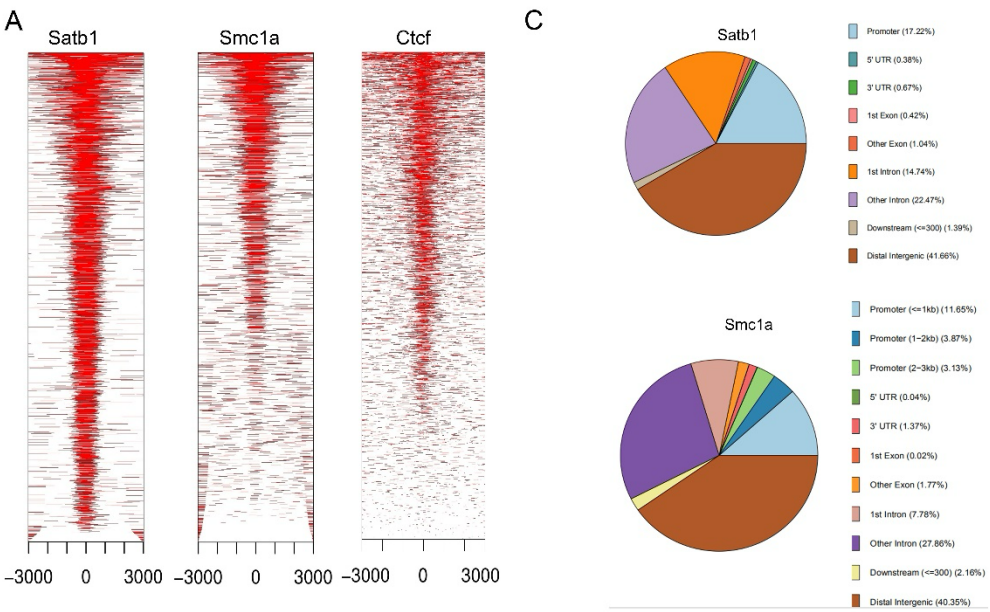


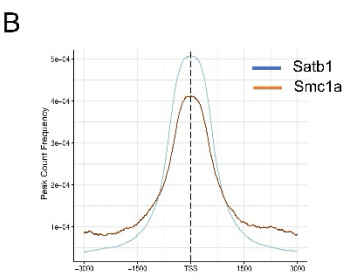
Figure 1.3.3: Dissection of interaction domains of Satb1 and Smc1a. A. Schematic of full length SATB1 and major domains of Satb1 which were cloned as HIS tagged ORFs in pTRIEX-3 Neo. **B.** Schematic representation of full-length and the functional domains of Smc1a. The individual domains, as depicted, were cloned into pGEX4T1 vector to tag them with glutathione S transferase (GST) as a fusion protein. **C.** Full-length SATB1 as well as PDZ/ULD and Cut/MAR domain (MD) + homeodomain (HD) were purified using Ni-NTA affinity purification. **D.** The individual domains of Smc1a were purified as GST-fusion proteins using glutathione

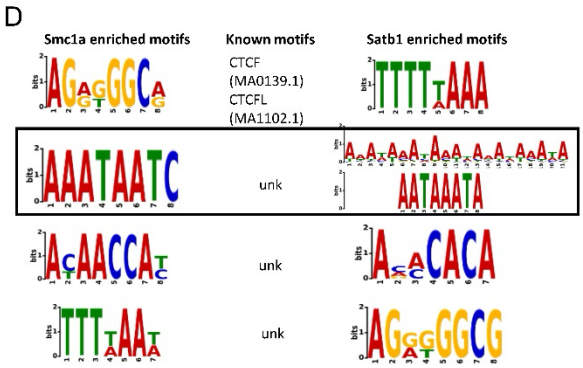
beads. Both panels **C** and **D** depict polyacrylamide gels (PAGE) stained with Coomassie blue. **E.** In vitro pull-down assay was carried out using full length HIS-Satb1 with individual domains of Smc1a. Ni-NTA pull down (PD) followed by western blotting of the PD with anti-GST antibody revealed the Satb1 interacting domains of Smc1a. **F.** Similar to **E**, a reverse assay was carried out in which full length Smc1a was incubated with the 2 arms of Satb1, individually. HIS-tag pull down followed by western blotting of Smc1a showed PDZ/ULD as uniquely interacting with Smc1a.

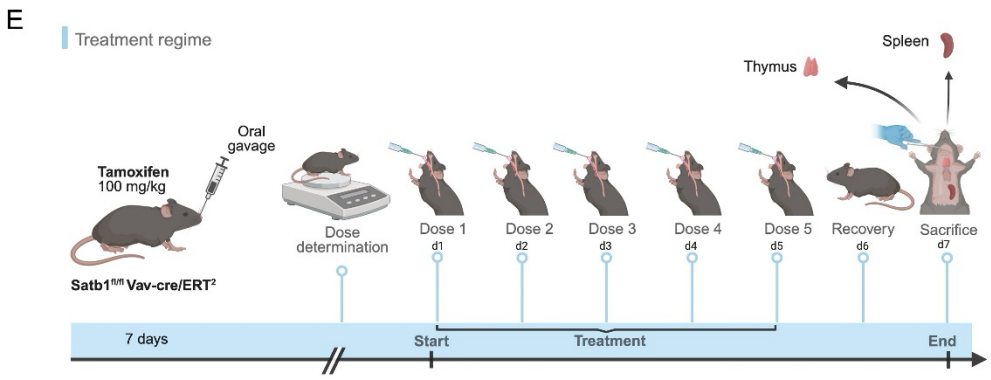
1.3.4 Generation of inducible Satb1 knockout in hematopoietic lineages

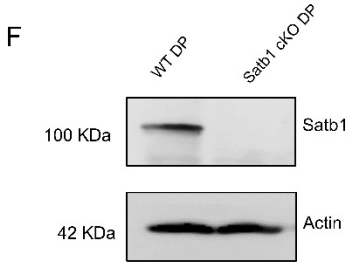
Satb1 being a strong nucleosome binding protein (Ghosh et al., 2019) possess ability to redirect its interacting TFs onto its genomic targets (Hosokawa et al., 2018; Notani et al., 2010). To ascertain a collaborative function of Satb1 with other architectural proteins and its interaction partner Smc1a, we analyzed chromatin occupancies of Satb1, Smc1a and Ctcf in DP thymocytes and found a strong correlation in their genomic occupancies at Satb1-bound regions. We observe that Smc1a displayed a broader occupancy profile than Ctcf in the tag-heatmaps, which conforms with that of Satb1 'broad' genomic binding (Fig 1.3.4A). Satb1 and Smc1a also show a similar average TSS occupancy. (Fig 1.3.4B). Although, both the proteins bind throughout the genome and share many annotated features (Fig 1.3.4C), it is interesting to note the higher occupancy of Satb1 at proximal promoters compared to Smc1a, alluding to its more direct regulatory role in the transcriptional control. Utilizing MEME-ChIP motif search, we also found that Satb1 consensus-like motif is enriched for Smc1a binding, although the top hits corresponded with Ctcf consensus site (Fig 1.3.4D). To delve into the functional role of Satb1 in T cell development, we utilized the inducible Vav-cre mediated knockout system (Vav-icre crossed with Satb1 fl/fl) to delete Satb1 from hematopoietic and T cell lineages upon tamoxifen treatment, and created a Satb1fl/fl-Vav-icre mouse line. Upon induction of ERT2 (Fig 1.3.4E), we observed more than 95% deletion of Satb1 from total thymocytes (Fg 1.3.4F). In line with previous studies (Feng et al., 2022; Kondo et al., 2016), we observed decrease in DN, CD4SP and CD8SP with a concomitant increase in DP thymocyte population (Fig 1.3.4G).


A
Satb1
Smc1a
Ctcf
−3000 0 3000 −3000 0 3000 −3000 0 3000
C
Satb1
Promoter (17.22%)
5' UTR (0.38%)
3' UTR (0.67%)
1st Exon (0.42%)
Other Exon (1.04%)
1st Intron (14.74%)
Other Intron (22.47%)
Downstream (<=300) (1.39%)
Distal Intergenic (41.66%)
Smc1a
Promoter (<=1kb) (11.65%)
Promoter (1~2kb) (3.87%)
Promoter (2~3kb) (3.13%)
5' UTR (0.04%)
3' UTR (1.37%)
1st Exon (0.02%)
Other Exon (1.77%)
1st Intron (7.78%)
Other Intron (27.86%)
Downstream (<=300) (2.16%)
Distal Intergenic (40.35%)


B
Satb1
Smc1a
Peak Count Frequency
−3000 −1500 TSS 1500 3000


D
Smc1a enriched motifs
Known motifs
Satb1 enriched motifs
CTCF (MA0139.1) CTCFL (MA1102.1)
unk
unk
unk


E
Treatment regime
Spleen
Thymus
Oral gavage
Tamoxifen 100 mg/kg
Satb1fl/fl Vav-cre/ERT2
Dose determination
Dose 1 d1
Dose 2 d2
Dose 3 d3
Dose 4 d4
Dose 5 d5
Recovery d6
Sacrifice d7
7 days
Start
Treatment
End


F
WT DP
Satb1 cKO DP
100 KDa
Satb1
42 KDa
Actin

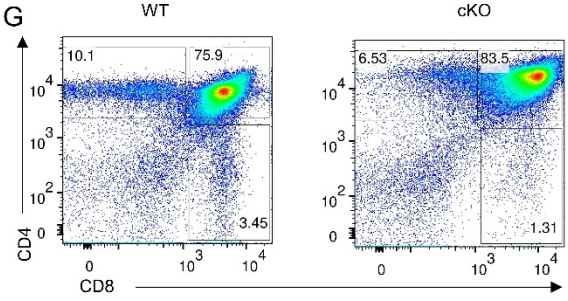

G
WT
cKO
10.1
75.9
3.45
6.53
83.5
1.31
CD4
CD8

Figure 1.3.4: Generation of inducible Satb1 knockout in hematopoietic lineages. A. Satb1 genomic co-occupancy with the cohesin core subunit Smc1a and Ctcf was assessed using publicly available ChIP seq datasets. Satb1 peaks were called and plotted as tag-density heatmaps. Smc1a and Ctcf datasets were overlayed with Satb1 peak-center and heatmaps were plotted, depicting their common genomic binding. **B.** Smc1a and Satb1 ChIP seq datasets were used to generate average TSS binding profiles, which were overlayed. **C.** ChIP peaks of Satb1 and Smc1a were annotated using R utility ChIPseeker to show various genomic features with which both the factors associate. **D.** Called peaks from Satb1 and Smc1a datasets were subjected to the MemeChIP suite and motif enrichment was carried out. Smc1a motif analysis revealed Satb1 consensus-like motif as one of the enriched motifs. **E.** Schematic of ERT2 mediated induction to knockout Satb1 in hematopoietic cells. Tamoxifen induces ERT2, promoting the expression of Cre protein in Vav positive cells to delete the floxed exon-2 region in Satb1 gene. **F.** Knockout efficiency of Satb1 in DP thymocytes upon tamoxifen treatment regime depicted in **E** via western blotting. n=3 **G.** Developing T cell populations in the thymus is shown via flow cytometry analysis. CD4 was stained with BV421 and Cd8 was stained with PE fluorophores. FACS plot is a representation of three biological replicates.

1.3.5 Satb1 deletion leads to partial repositioning of Smc1a on chromatin

Upon validating the knockout of Satb1 in DP thymocytes, we next asked whether depletion of Satb1 could lead to changes in chromatin occupancy of Smc1a. We performed Smc1a ChIP seq using SAtb1 WT and cKO DP thymocytes. We observed that overall TSS binding of Smc1a was slightly higher (Fig 1.3.5A) in cKO compared to WT samples. In contrast, the number of genes associated with Smc1a displayed a mild 9% decrease, with 1902 genes in WT Satb1 and 1738 genes in cKO Satb1 (Fig 1.3.5B). In the distinct genes' subsets, we then focused on the commonly enriched targets for Satb1 and Smc1a essential for T cell development. We found that at *CD8* super-enhancer (SE) cluster, the binding of Smc1a is diminished upon depletion of Satb1, also corresponding to the reduced chromatin accessibility. Whereas at *Cd4* promoter, Smc1a occupancy increased as a result of Satb1 cKO (Fig.1.3.5C). We validated the co-occupancy of Satb1 and Smc1a on the promoters of multiple immune-specific genes; Cd3e, Cd4, Cd8 and Il2ra via ChIP-qRT PCR (Fig1.3.5D). Overall, these data indicate that Smc1a genomic occupancy gets partially redirected in the absence Satb1 which might be essential for its spatial access to the chromatin.

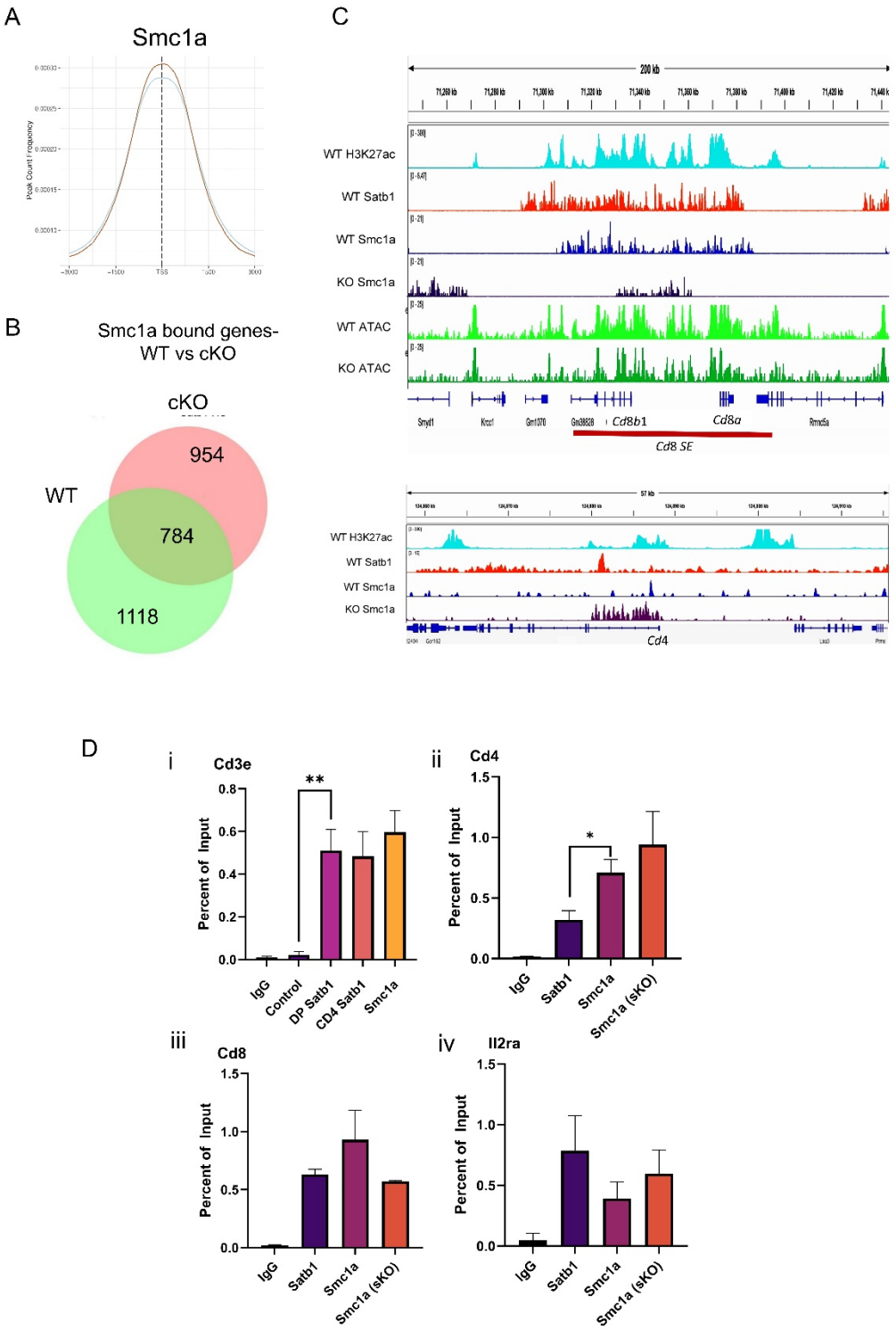


Figure 1.3.5: Satb1 deletion leads to partial repositioning of Smc1a on the chromatin. **A.** Smc1a ChIP-seq was performed in WT and Satb1 cKO DP thymocytes. Average TSS binding profiles of Smc1a occupancy is shown, indicating that Satb1 KO results in only limited modulation in Smc1a chromatin binding. **B.** Nearest genes to Smc1a enriched peaks were obtained in WT and Satb1 cKO conditions. The Venn diagram shows the frequency of common

and distinct genes bound by Smc1a in Satb1 WT and Satb1 cKO thymocytes. **C.** Genomic repositioning of Smc1a is depicted on Cd8 super enhancer (SE) loci and Cd4 promoter. At the Cd8 loci, Satb1 cKO leads to reduction in Smc1a occupancy. Whereas Satb1 restrains Smc1a binding on the Cd4 promoter. **D.** Chromatin co-occupancy of Satb1 and Smc1a, and Smc1a binding upon Satb1 KO were validated in DP thymocytes by ChIP-qPCR. Normalization was carried out using 1% Input. Three biological replicates were used to calculate statistical significance using Graphpad v10. ** $p<0.01$ * $p<0.05$.

1.3.6 Satb1 collaborates with Smc1a to mediate gene expression

To further understand the functional relation of Satb1 and Smc1a on thymocyte differentiation, we depleted Smc1a along with Satb1 using CRISPR/Cas9 lentiviral delivery system (Fig 1.3.6A). We observed ~ 70% knockdown efficiency for Smc1a compared to control transduced cells at the protein level (Fig. 1.3.6B). To assess the effect of the double depletion, RNA-seq was performed for control, Satb1 knockout (cKO) and Satb1-Smc1a double knockout (dKO) thymocytes. Interestingly, we observed a synergistic effect of Smc1a KO on the genes up- and down-regulated in Satb1 KO alone when compared to WT. (Fig 1.3.6C). We also confirmed using qRT-PCR that the transcript levels of some of their commonly enriched genes (Cd3e, Cd4 and Cd8a) decrease in cKO compared to WT samples, and are further downregulated, upon Smc1a-Satb1 double depletion (Fig 1.3.6D). To address their genomic co-regulation potential, we also performed a dual luciferase reporter assay using promoter of a previously known Satb1 target Il2 (Fig 1.3.6E) (Kumar et al., 2005). We observed that co-transfection of Satb1 and Smc1a compared to Satb1 alone and Smc1a alone resulted in increased normalized luciferase activity (Fig 1.3.6F). Overall, these data suggest that Smc1a and Satb1 could regulate immune specific genes in a co-dependent manner.

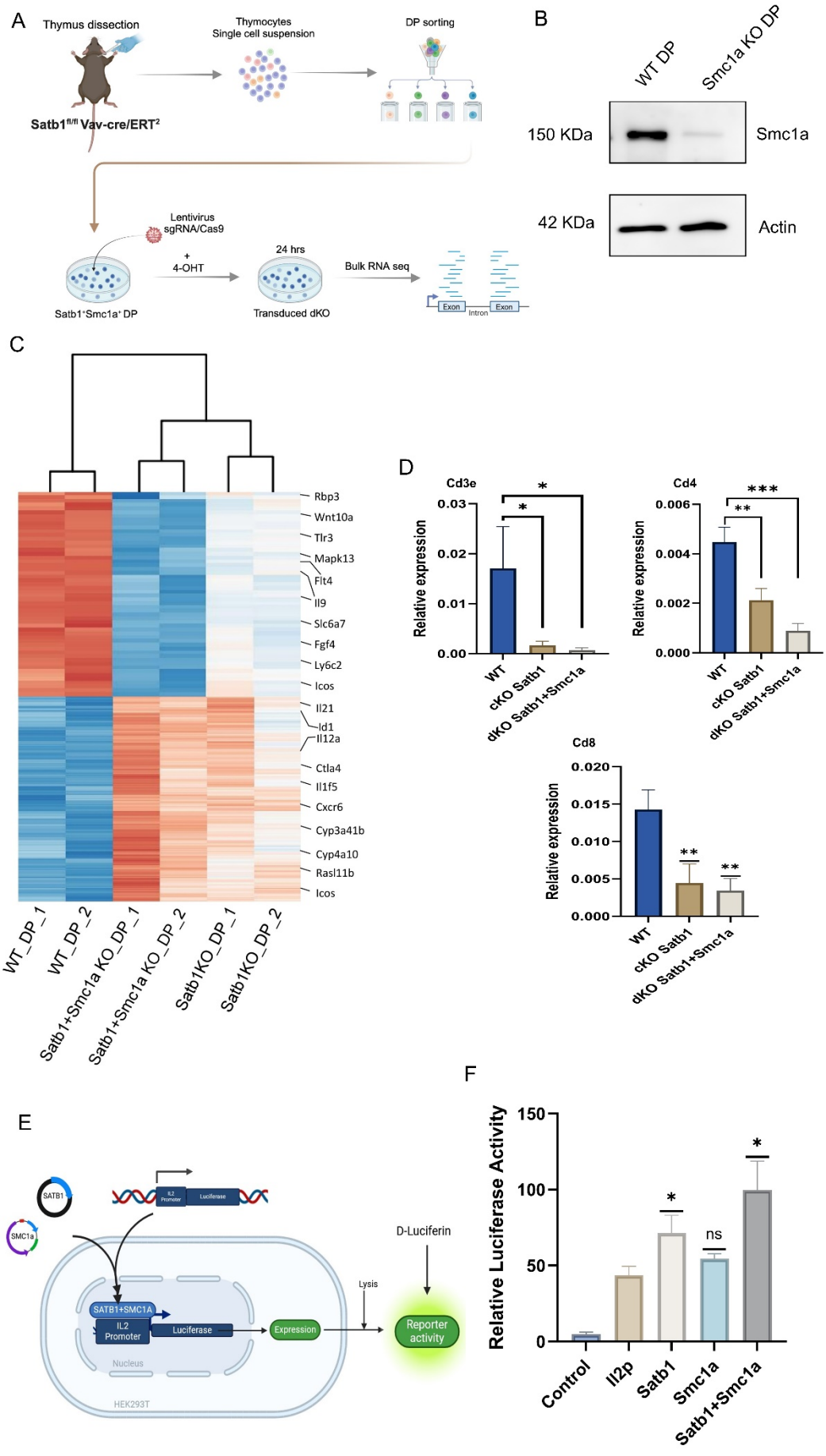

A
Thymus dissection
Satb1fl/fl Vav-cre/ERT2
Thymocytes
Single cell suspension
DP sorting
Lentivirus
sgRNA/Cas9
+
4-OHT
24 hrs
Bulk RNA seq
Satb1+Smc1a+ DP
Transduced dKO
Exon
Intron
Exon
B
WT DP
Smc1a KO DP
150 KDa
Smc1a
42 KDa
Actin
C
Rbp3
Wnt10a
Tlr3
Mapk13
Flt4
Il9
Slc6a7
Fgf4
Ly6c2
Icos
Il21
Id1
Il12a
Ctla4
Il1f5
Cxcr6
Cyp3a41b
Cyp4a10
Rasl11b
Icos
WT_DP_1
WT_DP_2
Satb1+Smc1a KO_DP_1
Satb1+Smc1a KO_DP_2
Satb1KO_DP_1
Satb1KO_DP_2
D
Cd3e
Relative expression
0.03
0.02
0.01
0.00
WT
cKO Satb1
dKO Satb1+Smc1a
Cd4
0.006
0.004
0.002
0.000
Cd8
0.020
0.015
0.010
0.005
0.000
E
SATB1
SMC1a
IL2 Promoter
Luciferase
SATB1+SMC1A
Expression
Lysis
D-Luciferin
Reporter activity
Nucleus
HEK293T
F
Relative Luciferase Activity
150
100
50
0
ns
Control
Il2p
Satb1
Smc1a
Satb1+Smc1a

Figure 1.3.6: Satb1 collaborates with Smc1a to mediate gene expression. A. Schematic representing the CRISPR strategy to carry out Satb1 Smc1a double knockout (dKO). **B.** Smc1a depletion was validated at the protein level compared to empty vector control. in DP thymocytes. **C.** Heatmap of top differentially expressing genes in WT and Satb1 and Smc1a dKO cells. The expression of these genes is also shown for Satb1 cKO as well, indicating that for multiple genes Satb1 and Smc1a could function in tandem to exert transcriptional regulation. **D.** RT-qPCR analysis of three target genes' expression, previously enriched for both Smc1a and Satb1 binding, in WT, Satb1 cKO and dKO conditions. All the three genes display a reduced expression in Satb1 KO, expression of which are further decreased, in the dKO condition. Expression of 18s rRNA was used to normalize gene expression, n=3. **E.** Schematic of luciferase assay utilized to assess the collaborative role of Satb1 and Smc1a in vitro. **F.** Relative luciferase activity estimation shows that Smc1a alone could not modulate the activity of Satb1 cognate target Il2 promoter. Whereas expression of Satb1 and Smc1a together increased its transcriptional activity significantly. Measurements of renilla (pTRK) activity was used to normalize the luciferase activity, n=4. Statistical significance was calculated using ANOVA in Graphpad v10. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

1.3.7 Satb1 maintains appropriate T cell activation during DP cell differentiation by modulating Cd3 expression

Upon successful V(D)J recombination of the TCR, DP thymocytes undergo positive and negative selection. Appropriate TCR signal is critical for their survival and transition to the SP stage (Bao et al., 2022; Choi et al., 2023). Previous studies suggested a role of Satb1 in regulating the differentiation, particularly to CD4 SP (Kondo et al., 2016) by regulating key receptor genes Cd4, Cd8 and Runx3, ThPOK and Foxp3 (Kakugawa et al., 2017). To understand the effect of Satb1 cKO on the positive selection of DP thymocytes, we monitored CD69, an early TCR-induced receptor (Fig 1.3.7A) (Yassai et al., 2002). We observed that CD69 surface expression was significantly reduced in Satb1 cKO DPs (17.4% $CD69^{+}$) compared to WT DP (11.3% $CD69^{+}$) (Fig 1.3.7A). We furthermore observed a stronger effect in CD69 levels in the differentiated CD4 (24.7% in WT and 4.8% in cKO) and CD8 (39.7% in WT and 19.7% in cKO) SP cells (Fig 1.3.7B). We also observed a significant reduction in TCRβ levels along with CD69 (Fig 1.3.7C), both together indicating a defect in positive selection (Yassai et al., 2002). In our data, transcript levels of Cd69 were also reduced significantly. Although no significant change in transcript levels at the *Tcrb* locus was observed in our RNA-seq datasets. To further probe into the activation defects, we monitored the *Cd3* locus, which

was earlier identified as one of Satb1 targets. CD3 complex is non-covalently associated with the cytoplasmic chains of the TCR heterodimer, and is crucial for facilitating the TCR signal to downstream molecules in the signalosome cascade (Bao et al., 2022; Klein et al., 2014; Smith-Garvin et al., 2009). From our flow cytometry data, we found that CD3e surface levels decreased dramatically in the positively selected CD4/CD8 SP populations upon Satb1 cKO (Fig 1.3.7D). Surprisingly, we did not observe a significant reduction in CD3 protein on the surface of Satb1 cKO DPs compared to WT despite significant reduction in its transcripts for Cd3e, CD3g and Cd3d chains (Fig 1.3.7E). Furthermore, to see if dysregulated protein expression is a consequence of Satb1 dependent genomic interactions, we utilized publicly available HiC data (Feng et al., 2022). Using HiC datasets for WT and Satb1 cKO in DP thymocytes, we find moderate reduction in intra-TAD interaction frequency at the *Cd3* locus upon Satb1 depletion (Fig 1.3.7F). We also identify a TAD-merging event at the nearest boundary ~200 kb downstream of the *Cd3* locus in Satb1 cKO compared to WT (Fig 1.3.7G). Collectively, these data indicate that Satb1 controls the expression of CD3 in positively selected CD4/CD8 SPs. Additionally, the regulatory function of Satb1 might involve chromatin looping at the *Cd3* locus which is partially disrupted at Satb1 cKO.

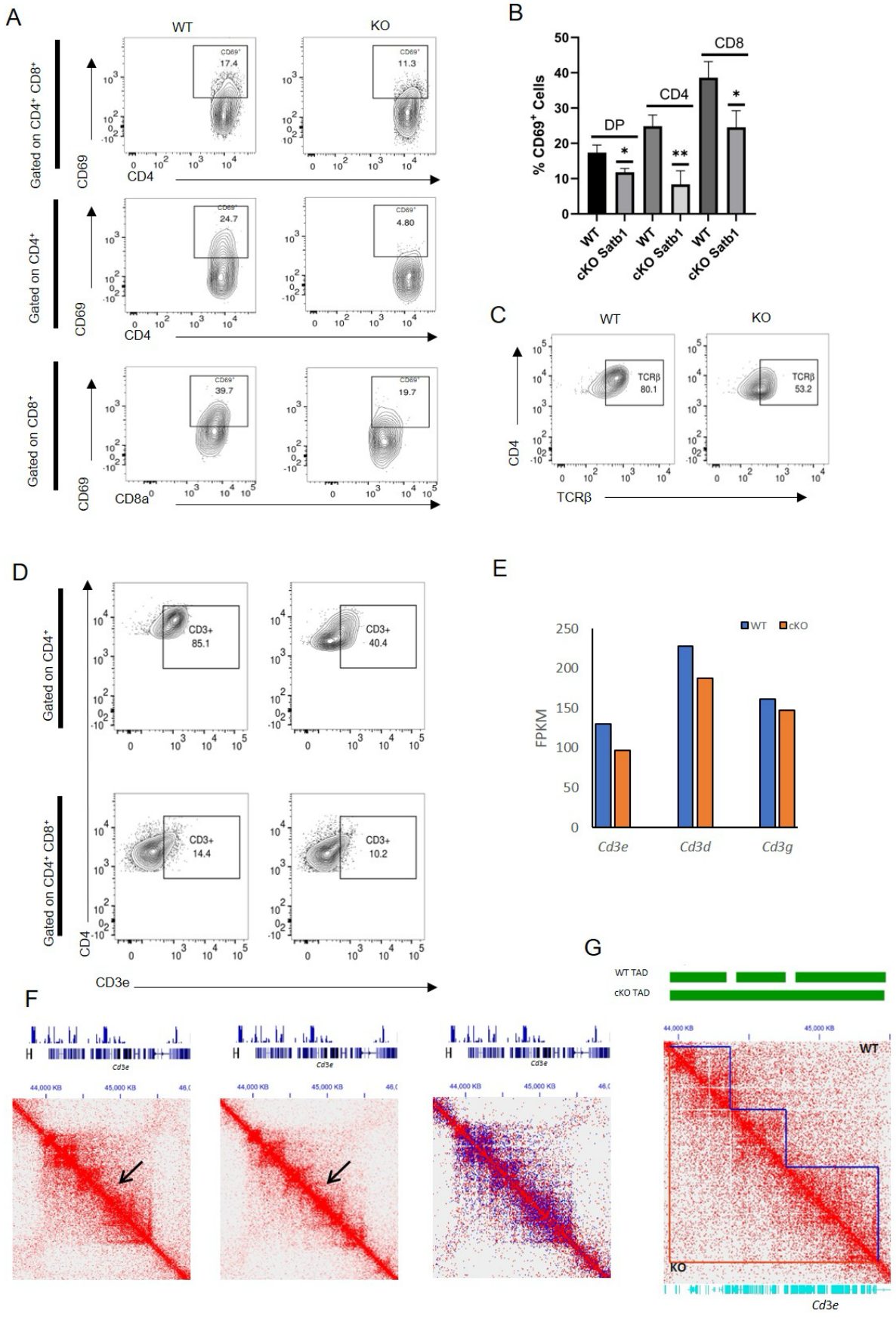

A
WT
KO
Gated on CD4+ CD8+
Gated on CD4+
Gated on CD8+
CD69
CD4
CD8a
17.4
11.3
24.7
4.80
39.7
19.7
B
% CD69+ Cells
DP
CD4
CD8
WT
cKO Satb1
C
TCRβ
80.1
53.2
D
CD3+
85.1
40.4
14.4
10.2
CD3e
E
FPKM
Cd3e
Cd3d
Cd3g
F
G
WT TAD
cKO TAD
44,000 KB
45,000 KB
WT
KO

Figure 1.3.7: Satb1 maintains appropriate T cell activation during DP cell differentiation by modulating Cd3 expression. A. Flow cytometry analysis of early T cell activation marker CD69 was performed in thymic T cells in WT and Satb1 cKO conditions. Cells were first gated for DP (CD4+CD8+), CD4 (CD4+CD8-) and CD8 (CD4-CD8+). The surface expression of CD69 is shown in these populations, depicting a stark reduction in expression as the cells undergo TCR-dependent differentiation. **B.** Quantitation of CD69+ cells in DP, CD4 and CD8 populations in WT and KO cells. While DP and CD8 show significant reduction of CD69 expression in Satb1 KO cells, CD4 SP shows a more dramatic reduction in CD69 levels. **C.** CD4 cells were gated and TCRβ surface expression was monitored. We observed a ~40% reduction in TCRβ surface levels in Satb1 KO thymocytes. Together CD69 and TCRβ reduction in Satb1 KO thymocytes indicated a defect in positive selection. **D.** CD3e surface levels in DP and CD4 SP thymocytes is shown. While DP does not show a significant change in CD3e expression, there is a dramatic reduction in CD3e levels in CD4 SP thymocytes, indicating a persistent requirement of Satb1 in positively selected SP thymocytes. **E.** Normalized expression of the three major Cd3 transcripts -Cd3e, Cd3g and Cd3d is plotted as FPKM values in WT and Satb1 KO conditions in DP thymocytes. **F.** Normalized HiC data from WT and Satb1 cKO in DP thymocytes was plotted using Juicebox. The *Cd3* locus shown depicts a decrease in genomic interactions within the TAD spanning the *Cd3* locus, which may affect CD3 expression in DP or later stages. **G.** TADs were called in WT and KO conditions. The TAD boundaries around the *Cd3* locus were plotted along with the interaction map of the same region. The loss of individual sub-TADs and thus formation of a merged large TAD is shown, depicting the requirement of Satb1 to maintain genomic conformation at the *Cd3* locus. FlowJo v10 was used for flow cytometry analysis (n=3) and MFI calculations. One mouse was used per experiment. Each was treated as a biological replicate. Representative images are shown here. Graphpad v10 was used for statistical analysis, n=3. * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

ChIP-seq analysis indicated that Satb1 binds at the super-enhancer cluster of Cd3 alongside Cohesin complex (Ing-Simmons et al., 2015), along with Ctcf (Fig 1.3.8A). Although, interestingly, we observed that at certain regions, Satb1 co-occupies the chromatin with Smc1a which is devoid of Ctcf. We thus tested whether Smc1a depletion along with Satb1 affected CD3 expression synergistically. Indeed, CD4/CD8 SP showed increased reduction in CD3e expression in dKO cells compared to Satb1 cKO alone or WT cells (Fig 1.3.8B and Fig 1.3.8C). These data demonstrate that Satb1 and Smc1a could cooperate to promote Cd3e expression in CD4/CD8 SP thymocytes.

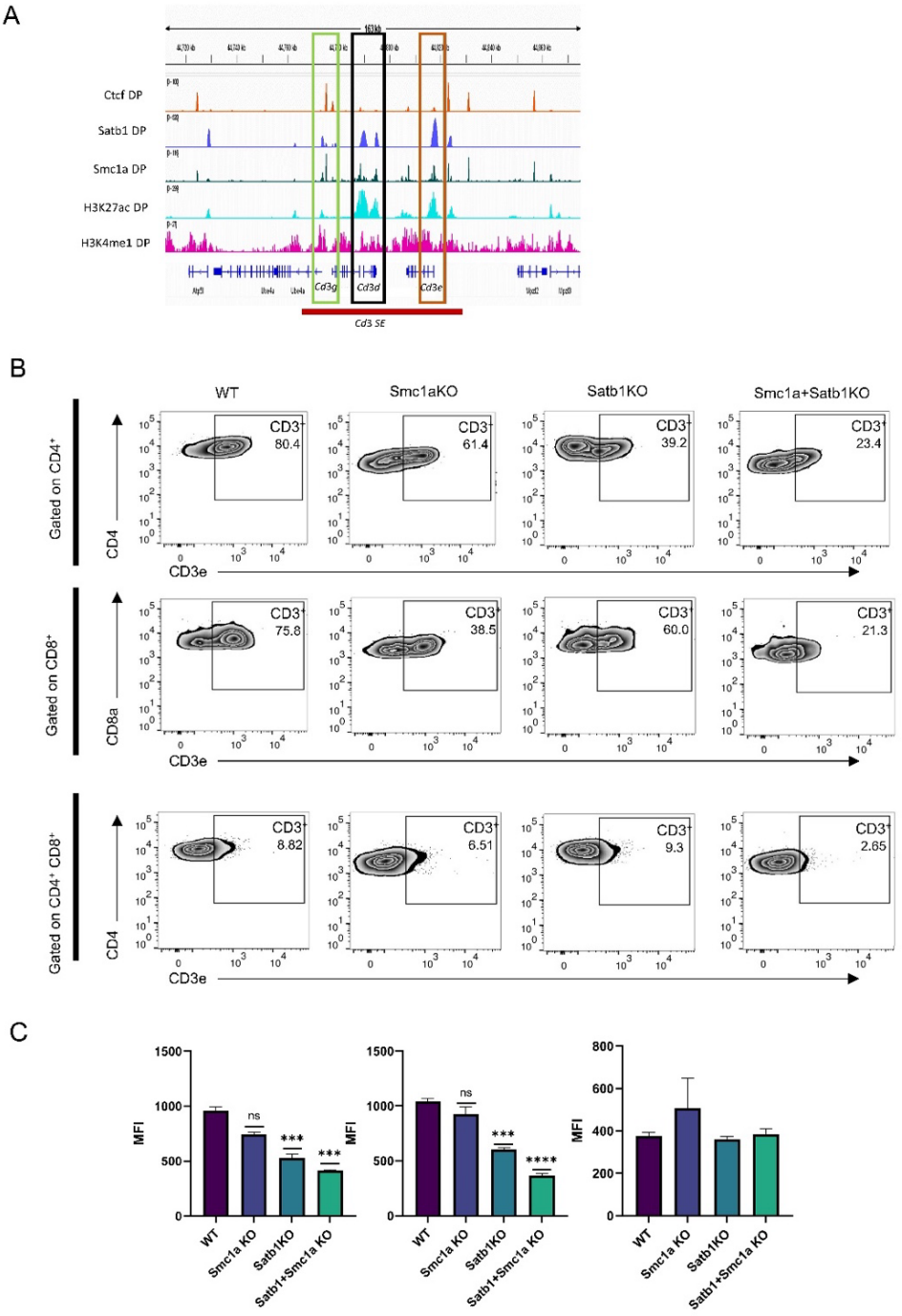


Figure 1.3.8: Satb1 and Smc1a could co-regulate CD3e expression in positively selected thymocytes. A. ChIP-seq occupancies of architectural factors Ctcf, Satb1, Smc1a, as well as histone modifications H3K27ac and H3K4me1, indicating the super enhancers (SE) were plotted at the *Cd3* locus. As shown Satb1 and Smc1a could together occupy Cd3e, excluding Ctcf. **B.** Flow cytometry analysis of DP, CD4 and CD8 thymocytes in WT, Satb1 KO, Smc1KO and Satb1 Smc1a dKO cells. As indicated, Satb1 KO as well as Smc1a KO results in reduction of CD3e expression. Whereas Smc1a+Satb1 dKO shows a further decrease in the CD3e surface levels in CD4 and CD8 SP cells, but not in DP thymocytes. **C.** Mean fluorescence intensity (MFI) was calculated for CD3e levels in all the conditions shown in **B**. FlowJo v10 was used for flow cytometry analysis (n=3) and MFI calculations. Representative images are shown here. Graphpad v10 was used for statistical analysis, n=3. * $p<0.05$ ** $p<0.01$ *** $p<0.001$.

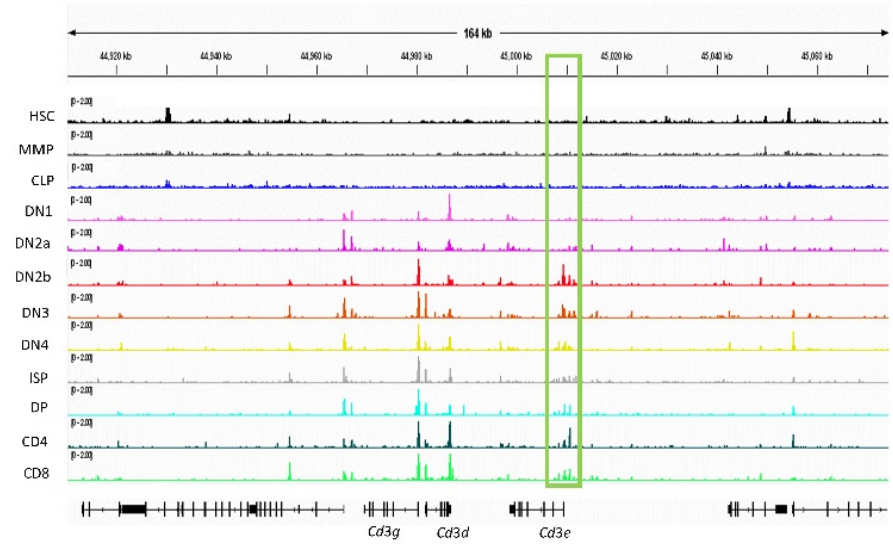
1.3.8 Satb1 is essential for chromatin accessibility at the *Cd3* locus during T cell development

It has been recently shown that during T cell development, there is a dramatic shift in the chromatin accessibility landscape (Hu et al., 2018) along with the global histone modifications as well as transcriptional paradigms (Zhang et al., 2012). Chromatin accessibility, as assayed by DNase I hypersensitive sequencing (DNase-seq) and assay for transposase-accessible chromatin (ATAC) sequencing revealed major shifts occurring at DN2b from ETP stages (Hu et al., 2018; Yoshida et al., 2019). Pioneer-factors TCF1 and BCL11b have been shown to modulate chromatin accessibility during different stages of T cell development (Gounari and Khazaie, 2022; Sidwell and Rothenberg, 2021). TCF1 could promote or preserve accessibility of chromatin during early T cell development (Emmanuel et al., 2018; Johnson et al., 2018). BCL11b, on the other hand has been demonstrated to open a closed chromatin state at the *TCRβ* locus which could explain its role in promoting the β-selection process (Sidwell and Rothenberg, 2021). We monitired the chromatin accessibility patterns at the *Cd3* locus. We observed that the *Cd3* locus is only partially accessible (at Cd3d) in the DN1 and DN2a stages, whereas from DN2b stage onwards, the all the *Cd3* genes gain accessibility for TF binding (Fig 1.3.9A)

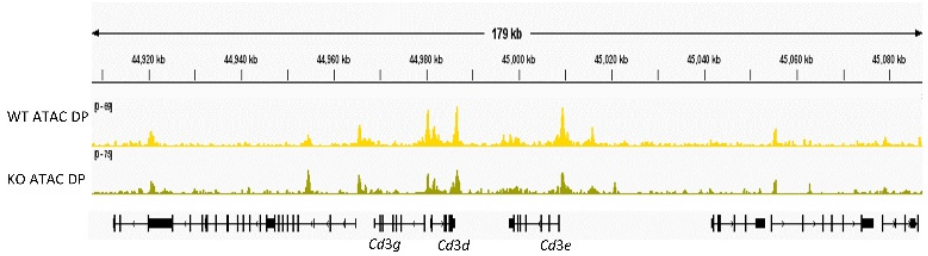
Satb1 also exhibits pioneer-like function by binding to both open- and closed-chromatin. Thus, we assessed whether Satb1 deletion could modulate chromatin accessibility at the *Cd3* locus. We utilized publicly available ATAC seq data from Satb1 WT and cKO DP cells (Feng et al., 2022). We saw an observable decrease in the ATAC-seq signal at the *Cd3* locus upon Satb1

knockout (Fig 1.3.9B). Although no change in histone modifications H3K4me3 and H3K27me3 was observed (Fig 1.3.9C). We also performed the transposase reactions in the 4 major stages namely DN, DP, CD4 and CD8 in WT and Satb1 cKO conditions followed by quantitative real-time PCR. We observed a modest decrease in accessibility at *CD3e* promoter in the DN and DP stages (Fig 1.3.9D). Whereas a strong reduction (~6fold) in amplification was observed in cKO CD4 SP when compared to WT CD4 SP (Fig 1.3.9E). Collectively, these results indicate that Satb1 could modulate the chromatin accessibility at the *Cd3* locus differently at different stages of T cell development, more dramatically at the positively selected SP stages.

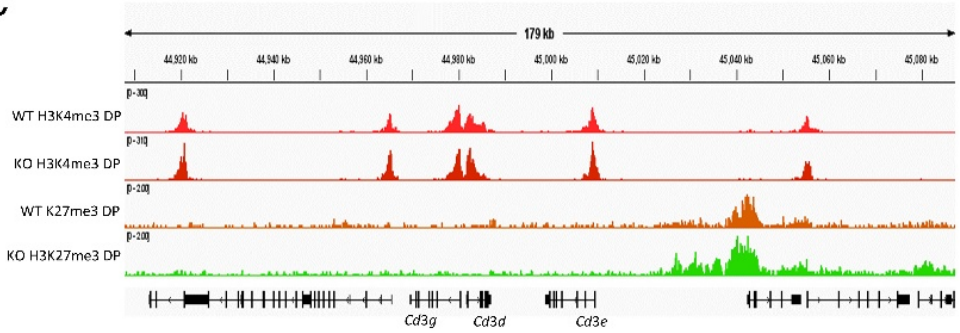
A


164 kb
HSC
MMP
CLP
DN1
DN2a
DN2b
DN3
DN4
ISP
DP
CD4
CD8
Cd3g
Cd3d
Cd3e

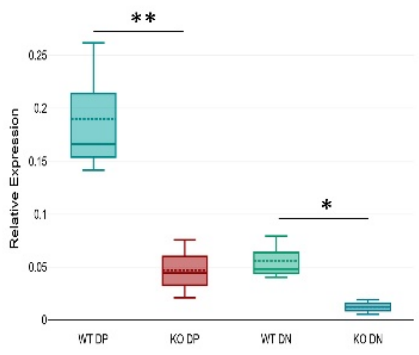
B


179 kb
WT ATAC DP
KO ATAC DP
Cd3g
Cd3d
Cd3e

C


179 kb
WT H3K4me3 DP
KO H3K4me3 DP
WT K27me3 DP
KO H3K27me3 DP
Cd3g
Cd3d
Cd3e

D


**
*
Relative Expression
WT DP
KO DP
WT DN
KO DN

E

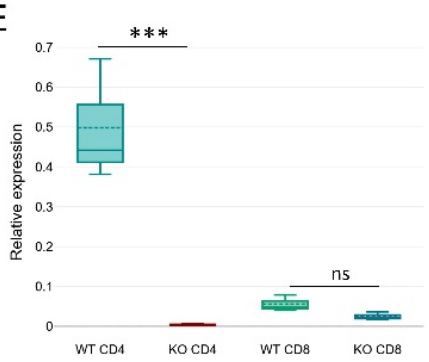

***
ns
Relative expression
WT CD4
KO CD4
WT CD8
KO CD8

Figure 1.3.9: Satb1 is essential for chromatin accessibility at the *Cd3* locus during T cell development. A. ATAC seq analysis of publicly available data of various pre-T, pro-T and committed T lineage cells, as indicated, was performed and plotted at the *Cd3* locus. As depicted, chromatin accessibility at the *Cd3* genes majorly increases during DN2a and DN2b stages, coinciding with the T lineage commitment, and the chromatin remains open in all the subsequent stages. **B.** WT and Satb1 cKO publicly available datasets for DP thymocytes were utilized to assess the effect of Satb1 in chromatin 'openness' at the Cd3 locus. **C.** ChIP seq data for histone modifications H3K4me3 and H3K27me3 in WT and Satb1 cKO conditions was analyzed and plotted at the *Cd3* locus. No change in their occupancy was observed as depicted in the plots. **D** and **E.** Assay for transposase-accessible chromatin (ATAC) was performed in WT and Satb1 KO sorted thymocytes, followed by quantitative PCR. ATAC-qPCR analysis revealed a significant reduction of *Cd3e* chromatin accessibility at the promoter region in Satb1 depleted DN, DP and CD4 populations compared to the WT cells. Duplicates were used for statistical significance using Graphpad v10, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

1.3.9 Satb1 is essential for maintaining CD4 T cell homeostasis in the periphery

Upon successful maturation in the thymus, CD4 and CD8 SP emigrate to the secondary lymphoid organs -spleen, lymph nodes, Peyer's patches and mucosal tissues (Ruterbusch et al., 2020). Both $CD4^+$ and $CD8^+$ retain a naïve status until they are exposed to pMHC complex while circulating. Upon the antigen presentation via pMHC-II or pMHC-I on the antigen presenting cells (APC), they undergo TCR activation and initiate a differentiation and cytotoxic transformation, respectively. Naïve $CD4^+$ T cells possess potential to differentiate into one of several subsets of effector T cells upon activation (Zhou et al., 2009). TCR in combination with CD3 activation leads to induction of various intracellular pathways leading to T cell proliferation and activation, priming them to select different differentiation dependent on the cytokine milieu (Luckheeram et al., 2012), TCR signal strength, type of APC and concentration of the antigen (Ashkar et al., 2000; Tao et al., 1997). The most relevant of APCs are dendritic cells (DCs) which possess enhanced capacity to stimulate naïve $CD4^+$ T cells, leading to their activation (Jenkins et al., 2001). We had previously shown that Satb1 plays a role in T cell differentiation toward T helper 2 (Th2) subset via regulating GATA3 expression and also mediate a cross-talk with Wnt-signaling (Notani et al., 2010). It was also shown that Satb1 deficiency leads to splenomegaly (Alvarez et al., 2000; Kondo et al., 2016). Therefore, we first assessed the effects of Satb1 cKO on the mature CD4 and CD8 populations in the spleen. Vav-cre mediated Satb1 cKO resulted in a dramatic reduction in both CD4 and CD8 populations, along with the

unusual occurrence of DP cells in the periphery (Fig 1.3.10A) which is in accordance with the recently published data (Kondo et al., 2016; Zelenka et al., 2022). We further observed that the naïve populations of CD4, as marked by CD62L ($CD62L^+$ $CD44^-$) are drastically diminished upon Satb1 cKO with a concomitant increase in the activated subset (Fig 1.3.10B). Upon identifying that the activation of CD4 T cells is disproportionate compared to WT counterparts, we further probed into the molecular role of Satb1 upon activation. To that end, we performed RNA-seq of activated and naïve $CD4^+$ T cells in WT and cKO conditions. We observed a dysregulation of activation specific, cytokine, and differentiation specific genes from the analysis (Fig 1.3.10C and 1.3.10D). We also performed network analysis on the downregulated genes upon Satb1 cKO in activated $CD4^+$ T cells. Top enriched factors corresponded to cytokine signaling and Wnt-mediated signaling networks (Fig 1.3.10E) suggesting a role of Satb1 in maintaining appropriate intracellular signaling upon T cell activation. We observed that IL4 receptor is also downregulated upon Satb1 cKO in activated $CD4^+$ T cells (Fig 1.3.10F) which could also partially explain the defects in Th2 differentiation. This is also corroborated by the reduction in ATAC seq peaks on the Il4ra promoter (Fig 1.3.10G). We furthermore checked whether Smc1a could augment Satb1 function in peripheral $CD4^+$ T cell activation. We again utilized the double KO strategy described previously in 1.3.6. We observed that upon dKO of Satb1 and Smc1a the naïve ($CD62^+$ $CD44^-$) subset did not reduce further compared to Satb1 cKO, but the increase in the effector ($CD62^-$ $CD44^+$) population was further exacerbated in the dKO condition (Fig 1.3.10H).

Together, these data suggest that Satb1 could maintain the naïve state of $CD4^+$ T cells by maintaining appropriate transcriptional program. Additionally, Smc1a augments Satb1 function by restricting aberrant increase in effector $CD4^+$ subsets.

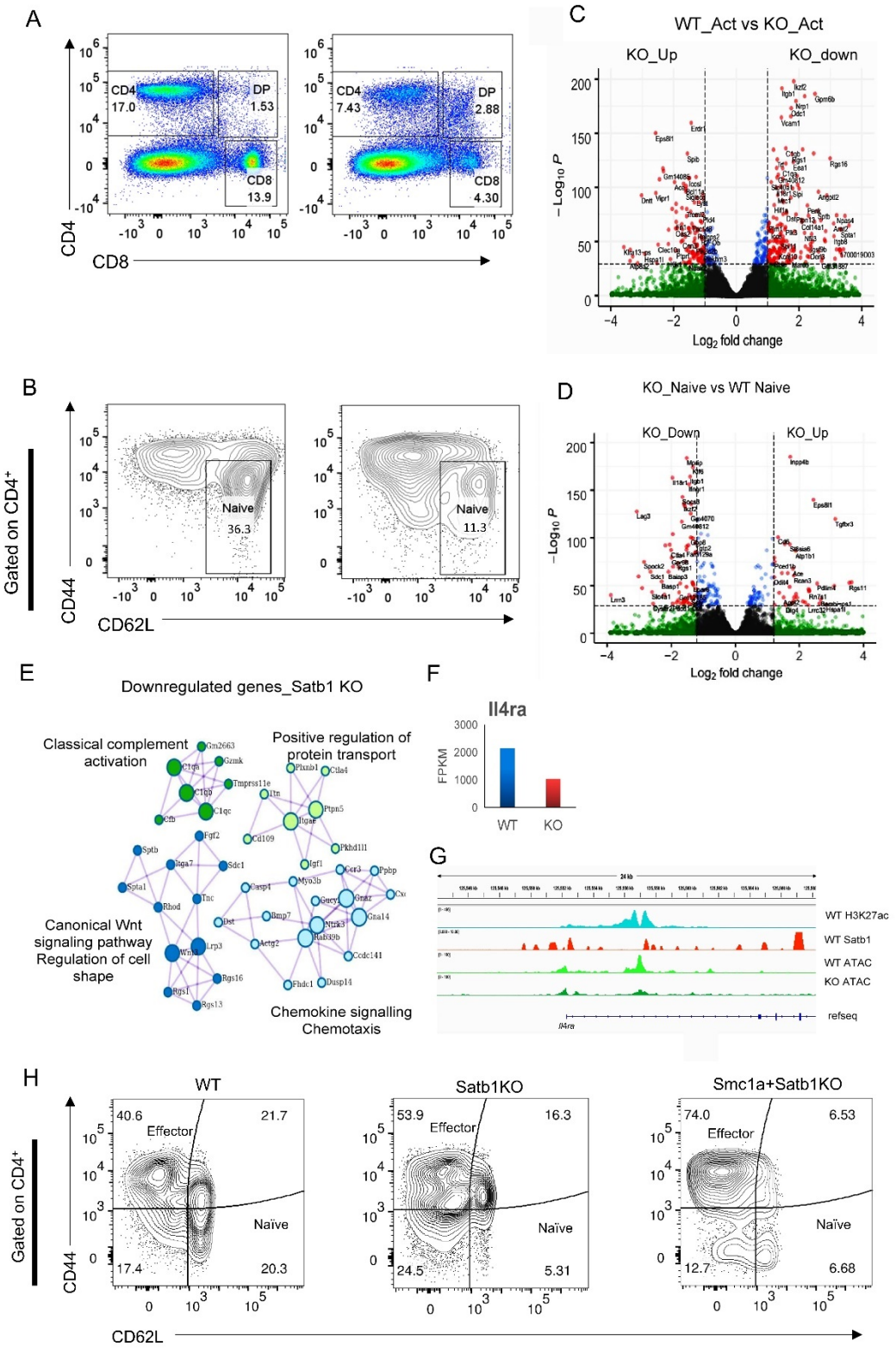

A
CD4
17.0
DP
1.53
CD8
13.9
CD4
7.43
DP
2.88
CD8
4.30
CD4
CD8
B
Gated on CD4+
Naive
36.3
Naive
11.3
CD44
CD62L
C
WT_Act vs KO_Act
KO_Up
KO_down
−Log10 P
Log2 fold change
D
KO_Naive vs WT Naive
KO_Down
KO_Up
−Log10 P
Log2 fold change
E
Downregulated genes_Satb1 KO
Classical complement activation
Positive regulation of protein transport
Canonical Wnt signaling pathway Regulation of cell shape
Chemokine signalling Chemotaxis
F
Il4ra
FPKM
WT
KO
G
WT H3K27ac
WT Satb1
WT ATAC
KO ATAC
refseq
H
WT
Satb1KO
Smc1a+Satb1KO
Gated on CD4+
40.6
Effector
21.7
Naïve
17.4
20.3
53.9
Effector
16.3
Naïve
24.5
5.31
74.0
Effector
6.53
Naïve
12.7
6.68
CD44
CD62L

Figure 1.3.10: Satb1 is essential for maintaining CD4 T cell homeostasis in the periphery. A. Flow cytometry analysis of splenic lymphocytes showing CD4, CD8 and DP populations in WT and Satb1 KO mice. As shown, a dramatic reduction in both mature CD4 and CD8 SP occurs as a result of Satb1 KO, as explained by their reduction in the thymus as well. On the contrary, DP cells showed an increase in the spleen. **B.** The CD4 SP cells were further gated into CD44 and CD62L subsets, which are markers for T cell activation. As shown, a reduction in the naïve CD4 cells was observed, indicating toward dysregulation in activation of CD4 T cells. **C** and **D.** Volcano plots from the RNA seq analysis depict a dysregulation in gene expression in both naïve and activated subsets of CD4 T cells upon SAtb1 depletion. **E.** The dataset for differential gene expression (DGE) in Satb1 KO versus WT activated CD4 cells was further subjected to network analysis. The most enriched pathways are depicted, suggesting the degree of involvement by Satb1 in regulating these signaling mechanisms during T cell activation. **F.** An example of Satb1 mediated priming, depicting Il4ra expression control during T cell activation. FPKM values of *Il2ra* gene is shown. **G.** ATAC seq analysis also confirms the importance of Satb1 in maintaining an 'open' chromatin state at the *Il4ra* promoter during T cell activation. **H.** Satb1 cooperation with Smc1a in regulating the naïve status of CD4 T cells is shown. Gated for CD4 SP subsets, the distribution and modulation of naïve and effector CD4 T cells is shown as a result of Satb1 KO and Satb1 Smc1a dKO compared to WT cells. Representative images of triplicates is shown for the flow cytometry data. Analysis of flow cytometry data was performed using FlowJO v10.

1.4 Discussion

As the thymic progenitors enter thymus and begin to receive Notch signaling, T lineage commitment begins and other lineage fates are progressively lost by the DN2b stage, which is coordinated by temporally controlled expression of different lineage-defining transcription factors (Hosokawa and Rothenberg, 2021). Satb1 is one such T lineage enriched transcription factor that has been studied in the context of genome organization and transcriptional control (Galande et al., 2007; Kumar et al., 2007; Yasui et al., 2002). Satb1 plays an essential role during T cell development and differentiation by modulating the expression of numerous T cell lineage genes and cytokines (Alvarez et al., 2000; Kondo et al., 2016). Satb1-null mice displays an arrest at the double positive stage (DP) of thymocyte development, accompanied by aberrant expression of IL2ra (Alvarez et al., 2000). During T cell development, Satb1 regulate the expression of CD4 specific genes such as Cd4, Zbtb7b (encoding ThPOK) (Kakugawa et al., 2017) and Gata3 (Notani et al., 2010). It has also been shown to promote Treg development from positively selected CD4 T cells by directly regulating Foxp3 expression via binding to its super-enhancer (SE). During T cell development, Satb1 expression strictly follows TCR activation and its expression varies in different stages of the development (Gottimukkala et al., 2016). During T cell development, to stringently control its expression, Satb1 utilizes alternative promoters which are different in their translations efficiencies and display a TCR-dependent switch in their usage (Patta et al., 2020). Although multiple facets of Satb1 function and its regulation in the T lineage context have been studied in the past three decades, its global function as an architectural protein is still not clear. In the current study, we have identified Satb1 specific genomic looping globally. Ctcf and to an extent cohesin have been studied extensively in the context of genome looping, TAD formation and maintenance of global chromatin architecture in various cell types (Stadhouders et al., 2019). Numerous genome-wide studies have investigated the correlation between genomic interactions and transcriptional readout (Beagan and Phillips-Cremins, 2020; Nora et al., 2012; Stadhouders et al., 2018). We observed that Satb1 also occupies majority of Ctcf as well as cohesin bound genomic sites in DP thymocytes. Ctcf is a not a cell-type dependent factor and is generally constitutively expressed (Dunn et al., 2003). It was earlier established that Ctcf knockout (KO) is is important to establish topologically associated domains in mouse embryonic stem (ES) cells (Chen et al., 2008; Nora et al., 2017). It is also demonstrated that Ctcf is not essential to maintain all chromatin domains once they are established, imparting minimal transcriptional dysregulation in its absence. Ctcf occupies around 60% of TAD boundaries in DP thymocytes. Using DP specific HiC analysis, we found that Satb1 occupies ~22% of TAD boundaries and associates with >60% of TAD associated genes, suggesting the potential of Satb1 to act as an architectural protein. Satb1 is specifically upregulated in T cells, primarily with the induction

of the TCR signal, may function in conjunction with Ctcf to add a more dynamic layer to the thymic 'loopscape' which is signal dependent.

In T cell development, multiple lineage-specifying transcription factors have been shown to function cooperatively. For instance, TCF1 has been recently shown to recruit HEB1 to its genomic sites to promote DP-specific program via modulating traditional enhancer (TE) and super-enhancer (SE) binding (Emmanuel et al., 2018). In thymocytes, Satb1 also interacts with β catenin in a Wnt-dependent manner (Notani et al., 2010) to cooperatively affect target gene expression. Satb1 interacts with multiple classes of proteins including histone modifiers such as HDACs (Kumar et al., 2005), p300 (Pavan Kumar et al., 2006), chromatin remodelers such as SWI/SNF, NuRD (Cai et al., 2003; Yasui et al., 2002), and transcription factors such as β-catenin in T cells. Our mass spectrometry data shows a strong interaction of Satb1 with factors involved in functions such as chromatin remodeling, Translation and chromatin modifications. One of the highly enriched interactors of Satb1 turned out to be Smc1a which is a core subunit of the cohesin complex. Cohesin complex has been previously implicated in condensation of chromatin, cell division and cell cycle. In its genomic roles, cohesin is responsible for the 'loop extrusion' of chromatin which is the leading mechanism for the formation of chromatin architecture in eukaryotes (Beagan and Phillips-Cremins, 2020). Loop extrusion mediated by cohesin is controlled by the presence of Ctcf at either end of the intervening DNA, interaction with which allows cohesin to release the DNA and stopping further extrusion. We show that Satb1 co-localizes with Smc1a in DP thymocytes. We further demonstrate that Satb1 gets immunoprecipitated along with Smc1a in thymocytes. At the genomic scale, Satb1 co-occupies ~70% of sites bound by Smc1a, sharing a majority of the regions with Ctcf as well. Satb1 and Smc1a co-bound genes are involved in numerous cellular processes, enriched for transcriptional co-regulation and Pol II activity. We also found that amongst Smc1a binding sites, one of the enriched motifs is Satb1 consensus-like sequence, demonstrating the preference commonality between the two proteins. In vitro, Satb1 and Smc1a retain the ability to interact directly, where we further dissected into the specific domains involved in the interaction.

The role of Satb1 T cell development has been explored previously using conditional knockout of Satb1 using Lck or CD4-cre, primarily in the late lineages of the development (Kondo et al., 2016; Yasuda et al., 2019) as well as Vav-cre specific knockout for all T cell lineages (Feng et al., 2022). To ascertain its functional roles in an inducible manner, we utilized inducible Vav-cre ERT2 mediated knockout of Satb1. We observed that Satb1 knockout leads to dramatic reduction CD4 and CD8 SP populations, along with a concomitant increase in DP subsets. Smc1a chromatin occupancy in DP thymocytes was altered upon Satb1 cKO, demonstrating a degree of Satb1 dependency by Smc1a for its chromatin homing.

We further demonstrate Smc1a and Satb1 in DP thymocytes coordinate co-activation of specific genes on which they display a co-occupancy, such as Cd3e, Cd4 and Cd8a. We also assessed the cooperation efficacy of Satb1 and Smc1a in vitro. Next, we demonstrated that Satb1 recruits Smc1a onto its target promoters to augment gene expression.

Satb1 deletion leads to differentiation defects from DP to CD4 and CD8 lineages, as well as in the development of Tregs (Kakugawa et al., 2017; Kondo et al., 2016), indicatively through dysregulated TCR signaling. We found that CD69 and TCR β expression are dramatically reduced in Satb1 knockout condition. Thus, Satb1 is essential for DP thymocytes to undergo positive selection. More strikingly, while DP shows a moderate yet significant reduction, the positively selected CD4/CD8 SP thymocytes show a much dramatic reduction in these molecules. This suggests a more prolonged role of Satb1 in maintaining the TCR-dependent activation state in mature SPs after their differentiation. We further demonstrated a pronounced role of Satb1 in regulating CD3 surface expression in positively selected SP thymocytes, which is a direct corollary of disrupted genomic interactions at the *Cd3* locus in Satb1 knockout conditions. This is plausibly the explanation of the poor responsiveness of DP thymocytes to TCR signal when presented with anti-CD3 and anti-CD28 signals (Kondo et al., 2016). Satb1 is a strong nucleosome binding protein (Ghosh et al., 2019). We also observed that Satb1 interacts with multiple histone proteins, indicative of its ability to bind closed chromatin. Therefore, we asked next whether Satb1 could modulate chromatin accessibility to regulate gene expression. We showed that at the *Cd3* locus, Satb1 is important for an open chromatin conformation in multiple stages of T cell development. We observed a more pronounced effect of Satb1 depletion in CD4 SP thymocytes. This becomes more plausible with the previous observations that Satb1 protein expression increases in CD4 thymocytes in response to higher TCR signaling. Thus, it could act as a feed-forward loop in maintenance of the signal response by the T cells.

The role of Satb1 in TCR activation is only recently been explored in the peripheral $CD4^+$ T cells (Gupta et al., 2022; Kuwabara et al., 2021) as well as in maintenance of naïve state in $CD8^+$ T cells (Nussing et al., 2022). We found that Satb1 deletion leads to significantly reduced $CD4^+$ and $CD8^+$ mature T cells. Furthermore, our data shows that naïve $CD4^+$ T cell population decreases dramatically with a concomitant rise in the activated population, which is ascribed to a shifted transcriptional status in these cells. Although the CD3 expression is still dependent positively on Satb1 in the CD4 subset, we think that the increased activation might be caused via cytokine dysregulation we observed.

In conclusion, we demonstrated that Satb1 can act as an architectural protein during T cell development in conjunction with Ctcf and Smc1a. Satb1 interacts with cohesin subunit Smc1a and could reposition its binding to chromatin based on the target sequence affinity, potentially affecting the looping mediated by the cohesin complex. Furthermore, Smc1a and Satb1 collaborate to drive transcriptional activation of a subset of the co-bound genes. We also demonstrated the importance of Satb1 in driving the expression of CD3e by maintaining the chromatin accessibility as well as genomic interactions at the *Cd3* locus. Finally, we show the importance of Satb1 in maintaining the activation versus naïve status of the $CD4^+$ T cells in the periphery.

1.5 References

Abramson, J., and Anderson, G. (2017). Thymic Epithelial Cells. Annu Rev Immunol *35*, 85-118.

Adoro, S., Erman, B., Sarafova, S.D., Van Laethem, F., Park, J.H., Feigenbaum, L., and Singer, A. (2008). Targeting CD4 coreceptor expression to postselection thymocytes reveals that CD4/CD8 lineage choice is neither error-prone nor stochastic. J Immunol *181*, 6975-6983.

Ahlfors, H., Limaye, A., Elo, L.L., Tuomela, S., Burute, M., Gottimukkala, K.V., Notani, D., Rasool, O., Galande, S., and Lahesmaa, R. (2010). SATB1 dictates expression of multiple genes including IL-5 involved in human T helper cell differentiation. Blood *116*, 1443-1453.

Alam, S.M., Travers, P.J., Wung, J.L., Nasholds, W., Redpath, S., Jameson, S.C., and Gascoigne, N.R. (1996). T-cell-receptor affinity and thymocyte positive selection. Nature *381*, 616-620.

Alipour, E., and Marko, J.F. (2012). Self-organization of domain structures by DNA-loop-extruding enzymes. Nucleic Acids Res *40*, 11202-11212.

Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev *14*, 521-535.

Ashkar, S., Weber, G.F., Panoutsakopoulou, V., Sanchirico, M.E., Jansson, M., Zawaideh, S., Rittling, S.R., Denhardt, D.T., Glimcher, M.J., and Cantor, H. (2000). Eta-1 (osteopontin): an early component of type-1 (cell-mediated) immunity. Science *287*, 860-864.

Bain, G., Engel, I., Robanus Maandag, E.C., te Riele, H.P., Voland, J.R., Sharp, L.L., Chun, J., Huey, B., Pinkel, D., and Murre, C. (1997). E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol *17*, 4782-4791.

Bao, X., Qin, Y., Lu, L., and Zheng, M. (2022). Transcriptional Regulation of Early T-Lymphocyte Development in Thymus. Front Immunol *13*, 884569.

Beagan, J.A., and Phillips-Cremins, J.E. (2020). On the existence and functionality of topologically associating domains. Nat Genet *52*, 8-16.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A.*, et al.* (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572 e524.

Bouillet, P., Purton, J.F., Godfrey, D.I., Zhang, L.C., Coultas, L., Puthalakath, H., Pellegrini, M., Cory, S., Adams, J.M., and Strasser, A. (2002). BH3-only Bcl-2 family member Bim is required for apoptosis of autoreactive thymocytes. Nature *415*, 922-926.

Braunstein, M., and Anderson, M.K. (2012). HEB in the spotlight: Transcriptional regulation of T-cell specification, commitment, and developmental plasticity. Clin Dev Immunol *2012*, 678705.

Burute, M., Gottimukkala, K., and Galande, S. (2012). Chromatin organizer SATB1 is an important determinant of T-cell differentiation. Immunol Cell Biol *90*, 852-859.

Cai, S., Han, H.J., and Kohwi-Shigematsu, T. (2003). Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nat Genet *34*, 42-51.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.

Cantrell, D. (1996). T cell antigen receptor signal transduction pathways. Annu Rev Immunol *14*, 259-274.

Carpenter, A.C., and Bosselut, R. (2010). Decision checkpoints in the thymus. Nat Immunol *11*, 666-673.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J.*, et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Choi, S., Lee, J., Hatzihristidis, T., Gaud, G., Dutta, A., Arya, A., Clubb, L.M., Stamos, D.B., Markovics, A., Mikecz, K.*, et al.* (2023). THEMIS increases TCR signaling in CD4(+)CD8(+) thymocytes by inhibiting the activity of the tyrosine phosphatase SHP1. Sci Signal *16*, eade1274.

Ciofani, M., and Zuniga-Pflucker, J.C. (2006). A survival guide to early T cell development. Immunol Res *34*, 117-132.

Cuadrado, A., and Losada, A. (2020). Specialized functions of cohesins STAG1 and STAG2 in 3D genome architecture. Curr Opin Genet Dev *61*, 9-16.

Dave, V.P., Allman, D., Keefe, R., Hardy, R.R., and Kappes, D.J. (1998). HD mice: a novel mouse mutant with a specific defect in the generation of CD4(+) T cells. Proc Natl Acad Sci U S A *95*, 8187-8192.

Davidson, I.F., and Peters, J.M. (2021). Genome folding through loop extrusion by SMC complexes. Nat Rev Mol Cell Biol *22*, 445-464.

Davis, C.B., Killeen, N., Crooks, M.E., Raulet, D., and Littman, D.R. (1993). Evidence for a stochastic mechanism in the differentiation of mature subsets of T lymphocytes. Cell *73*, 237-247.

Del Real, M.M., and Rothenberg, E.V. (2013). Architecture of a lymphomyeloid developmental switch controlled by PU.1, Notch and Gata3. Development *140*, 1207-1219.

Di Giammartino, D.C., Kloetgen, A., Polyzos, A., Liu, Y., Kim, D., Murphy, D., Abuhashem, A., Cavaliere, P., Aronson, B., Shah, V.*, et al.* (2019). KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. Nat Cell Biol *21*, 1179-1190.

Dias, S., Mansson, R., Gurbuxani, S., Sigvardsson, M., and Kee, B.L. (2008). E2A proteins promote development of lymphoid-primed multipotent progenitors. Immunity *29*, 217-227.

Dickinson, L.A., Dickinson, C.D., and Kohwi-Shigematsu, T. (1997). An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. J Biol Chem *272*, 11463-11470.

Dickinson, L.A., Joh, T., Kohwi, Y., and Kohwi-Shigematsu, T. (1992). A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. Cell *70*, 631-645.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schujiers, J., Lee, T.I., Zhao, K.*, et al.* (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374-387.

Dunn, K.L., Zhao, H., and Davie, J.R. (2003). The insulator binding protein CTCF associates with the nuclear matrix. Exp Cell Res *288*, 218-223.

Dutta, A., Zhao, B., and Love, P.E. (2021). New insights into TCR beta-selection. Trends Immunol *42*, 735-750.

Egawa, T., and Littman, D.R. (2008). ThPOK acts late in specification of the helper T cell lineage and suppresses Runx-mediated commitment to the cytotoxic T cell lineage. Nat Immunol *9*, 1131-1139.

Emmanuel, A.O., Arnovitz, S., Haghi, L., Mathur, P.S., Mondal, S., Quandt, J., Okoreeh, M.K., Maienschein-Cline, M., Khazaie, K., Dose, M.*, et al.* (2018). TCF-1 and HEB cooperate to establish the epigenetic and transcription profiles of CD4(+)CD8(+) thymocytes. Nat Immunol *19*, 1366-1378.

Feng, D., Chen, Y., Dai, R., Bian, S., Xue, W., Zhu, Y., Li, Z., Yang, Y., Zhang, Y., Zhang, J.*, et al.* (2022). Chromatin organizer SATB1 controls the cell identity of CD4(+) CD8(+) double-positive thymocytes by regulating the activity of super-enhancers. Nat Commun *13*, 5554.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.

Galande, S., Dickinson, L.A., Mian, I.S., Sikorska, M., and Kohwi-Shigematsu, T. (2001). SATB1 cleavage by caspase 6 disrupts PDZ domain-mediated dimerization, causing detachment from chromatin early in T-cell apoptosis. Mol Cell Biol *21*, 5591-5604.

Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev *17*, 408-414.

Gao, Y., Zamisch, M., Vacchio, M., Chopp, L., Ciucci, T., Paine, E.L., Lyons, G.C., Nie, J., Xiao, Q., Zvezdova, E.*, et al.* (2022). NuRD complex recruitment to Thpok mediates CD4(+) T cell lineage differentiation. Sci Immunol *7*, eabn5917.

Garcia-Ojeda, M.E., Klein Wolterink, R.G., Lemaitre, F., Richard-Le Goff, O., Hasan, M., Hendriks, R.W., Cumano, A., and Di Santo, J.P. (2013). GATA-3 promotes T-cell specification by repressing B-cell potential in pro-T cells in mice. Blood *121*, 1749-1759.

Gascoigne, N.R., Casas, J., Brzostek, J., and Rybakin, V. (2011). Initiation of TCR phosphorylation and signal transduction. Front Immunol *2*, 72.

Germain, R.N. (2002). T-cell development and the CD4-CD8 lineage decision. Nat Rev Immunol *2*, 309-322.

Ghosh, R.P., Shi, Q., Yang, L., Reddick, M.P., Nikitina, T., Zhurkin, V.B., Fordyce, P., Stasevich, T.J., Chang, H.Y., Greenleaf, W.J.*, et al.* (2019). Satb1 integrates DNA binding

site geometry and torsional stress to differentially target nucleosome-dense regions. Nat Commun *10*, 3221.

Goolam, M., and Zernicka-Goetz, M. (2017). The chromatin modifier Satb1 regulates cell fate through Fgf signalling in the early mouse embryo. Development *144*, 1450-1461.

Gottimukkala, K.P., Jangid, R., Patta, I., Sultana, D.A., Sharma, A., Misra-Sen, J., and Galande, S. (2016). Regulation of SATB1 during thymocyte development by TCR signaling. Mol Immunol *77*, 34-43.

Gounari, F., and Khazaie, K. (2022). TCF-1: a maverick in T cell development and function. Nat Immunol *23*, 671-678.

Guirado, M., de Aos, I., Orta, T., Rivas, L., Terhorst, C., Zubiaur, M., and Sancho, J. (2002). Phosphorylation of the N-terminal and C-terminal CD3-epsilon-ITAM tyrosines is differentially regulated in T cells. Biochem Biophys Res Commun *291*, 574-581.

Gupta, P.K., Allocco, J.B., Fraipont, J.M., McKeague, M.L., Wang, P., Andrade, M.S., McIntosh, C., Chen, L., Wang, Y., Li, Y*., et al.* (2022). Reduced Satb1 expression predisposes CD4(+) T conventional cells to Treg suppression and promotes transplant survival. Proc Natl Acad Sci U S A *119*, e2205062119.

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

He, X., He, X., Dave, V.P., Zhang, Y., Hua, X., Nicolas, E., Xu, W., Roe, B.A., and Kappes, D.J. (2005). The zinc finger transcription factor Th-POK regulates CD4 versus CD8 T-cell lineage commitment. Nature *433*, 826-833.

Heath, H., Ribeiro de Almeida, C., Sleutels, F., Dingjan, G., van de Nobelen, S., Jonkers, I., Ling, K.W., Gribnau, J., Renkawitz, R., Grosveld, F*., et al.* (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. EMBO J *27*, 2839-2850.

Heidinger-Pauli, J.M., Unal, E., Guacci, V., and Koshland, D. (2008). The kleisin subunit of cohesin dictates damage-induced cohesion. Mol Cell *31*, 47-56.

Hernandez-Hoyos, G., Anderson, M.K., Wang, C., Rothenberg, E.V., and Alberola-Ila, J. (2003). GATA-3 expression is controlled by TCR signals and regulates CD4/CD8 differentiation. Immunity *19*, 83-94.

Hoffman, E.S., Passoni, L., Crompton, T., Leu, T.M., Schatz, D.G., Koff, A., Owen, M.J., and Hayday, A.C. (1996). Productive T-cell receptor beta-chain gene rearrangement: coincident regulation of cell cycle and clonality during development in vivo. Genes Dev *10*, 948-962.

Hosokawa, H., and Rothenberg, E.V. (2021). How transcription factors drive choice of the T cell fate. Nat Rev Immunol *21*, 162-176.

Hosokawa, H., Ungerback, J., Wang, X., Matsumoto, M., Nakayama, K.I., Cohen, S.M., Tanaka, T., and Rothenberg, E.V. (2018). Transcription Factor PU.1 Represses and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription Factor Binding. Immunity *49*, 782.

Hozumi, K., Abe, N., Chiba, S., Hirai, H., and Habu, S. (2003). Active form of Notch members can enforce T lymphopoiesis on lymphoid progenitors in the monolayer culture specific for B cell development. J Immunol *170*, 4973-4979.

Hozumi, K., Mailhos, C., Negishi, N., Hirano, K., Yahata, T., Ando, K., Zuklys, S., Hollander, G.A., Shima, D.T., and Habu, S. (2008). Delta-like 4 is indispensable in thymic environment specific for T cell development. J Exp Med *205*, 2507-2513.

Hu, G., Cui, K., Fang, D., Hirose, S., Wang, X., Wangsa, D., Jin, W., Ried, T., Liu, P., Zhu, J., *et al.* (2018). Transformation of Accessible Chromatin and 3D Nucleome Underlies Lineage Commitment of Early T Cells. Immunity *48*, 227-242 e228.

Hwang, J.R., Byeon, Y., Kim, D., and Park, S.G. (2020). Recent insights of T cell receptor-mediated signaling pathways for T cell activation and development. Exp Mol Med *52*, 750-761.

Ikawa, T., Fujimoto, S., Kawamoto, H., Katsura, Y., and Yokota, Y. (2001). Commitment to natural killer cells requires the helix-loop-helix inhibitor Id2. Proc Natl Acad Sci U S A *98*, 5164-5169.

Ikawa, T., Hirose, S., Masuda, K., Kakugawa, K., Satoh, R., Shibano-Satoh, A., Kominami, R., Katsura, Y., and Kawamoto, H. (2010). An essential developmental checkpoint for production of the T cell lineage. Science *329*, 93-96.

Ikawa, T., Kawamoto, H., Goldrath, A.W., and Murre, C. (2006). E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment. J Exp Med *203*, 1329-1342.

Ing-Simmons, E., Seitan, V.C., Faure, A.J., Flicek, P., Carroll, T., Dekker, J., Fisher, A.G., Lenhard, B., and Merkenschlager, M. (2015). Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. Genome Res *25*, 504-513.

Iwashima, M. (2003). Kinetic perspectives of T cell antigen receptor signaling. A two-tier model for T cell full activation. Immunol Rev *191*, 196-210.

Iwashima, M., Irving, B.A., van Oers, N.S., Chan, A.C., and Weiss, A. (1994). Sequential interactions of the TCR with two distinct cytoplasmic tyrosine kinases. Science *263*, 1136-1139.

Jenkins, M.K., Khoruts, A., Ingulli, E., Mueller, D.L., McSorley, S.J., Reinhardt, R.L., Itano, A., and Pape, K.A. (2001). In vivo activation of antigen-specific CD4 T cells. Annu Rev Immunol *19*, 23-45.

Ji, H., Ehrlich, L.I., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M.J., Irizarry, R.A., Kim, K.*, et al.* (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature *467*, 338-342.

Johanson, T.M., Lun, A.T.L., Coughlan, H.D., Tan, T., Smyth, G.K., Nutt, S.L., and Allan, R.S. (2018). Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. Nat Immunol *19*, 1257-1264.

Johnson, J.L., Georgakilas, G., Petrovic, J., Kurachi, M., Cai, S., Harly, C., Pear, W.S., Bhandoola, A., Wherry, E.J., and Vahedi, G. (2018). Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells. Immunity *48*, 243-257 e210.

Jones, M.E., and Zhuang, Y. (2011). Stage-specific functions of E-proteins at the beta-selection and T-cell receptor checkpoints during thymocyte development. Immunol Res *49*, 202-215.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kakugawa, K., Kojo, S., Tanaka, H., Seo, W., Endo, T.A., Kitagawa, Y., Muroi, S., Tenno, M., Yasmin, N., Kohwi, Y.*, et al.* (2017). Essential Roles of SATB1 in Specifying T Lymphocyte Subsets. Cell Rep *19*, 1176-1188.

Kearse, K.P., Takahama, Y., Punt, J.A., Sharrow, S.O., and Singer, A. (1995). Early molecular events induced by T cell receptor (TCR) signaling in immature CD4+ CD8+ thymocytes: increased synthesis of TCR-alpha protein is an early response to TCR signaling

that compensates for TCR-alpha instability, improves TCR assembly, and parallels other indicators of positive selection. J Exp Med *181*, 193-202.

Kim, E., Barth, R., and Dekker, C. (2023). Looping the Genome with SMC Complexes. Annu Rev Biochem *92*, 15-41.

Kitagawa, Y., Ohkura, N., Kidani, Y., Vandenbon, A., Hirota, K., Kawakami, R., Yasuda, K., Motooka, D., Nakamura, S., Kondo, M.*, et al.* (2017). Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. Nat Immunol *18*, 173-183.

Klein, L., Kyewski, B., Allen, P.M., and Hogquist, K.A. (2014). Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). Nat Rev Immunol *14*, 377-391.

Klenova, E.M., Nicolas, R.H., Paterson, H.F., Carne, A.F., Heath, C.M., Goodwin, G.H., Neiman, P.E., and Lobanenkov, V.V. (1993). CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. Mol Cell Biol *13*, 7612-7624.

Koch, U., Fiorini, E., Benedito, R., Besseyrias, V., Schuster-Gossler, K., Pierres, M., Manley, N.R., Duarte, A., Macdonald, H.R., and Radtke, F. (2008). Delta-like 4 is the essential, nonredundant ligand for Notch1 during thymic T cell lineage commitment. J Exp Med *205*, 2515-2523.

Kominami, R. (2012). Role of the transcription factor Bcl11b in development and lymphomagenesis. Proc Jpn Acad Ser B Phys Biol Sci *88*, 72-87.

Kondo, M., Tanaka, Y., Kuwabara, T., Naito, T., Kohwi-Shigematsu, T., and Watanabe, A. (2016). SATB1 Plays a Critical Role in Establishment of Immune Tolerance. J Immunol *196*, 563-572.

Krangel, M.S. (2009). Mechanics of T cell receptor gene rearrangement. Curr Opin Immunol *21*, 133-139.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galande, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol *9*, 45-56.

Kumar, P.P., Purbey, P.K., Ravi, D.S., Mitra, D., and Galande, S. (2005). Displacement of SATB1-bound histone deacetylase 1 corepressor by the human immunodeficiency virus type 1 transactivator induces expression of interleukin-2 and its receptor in T cells. Mol Cell Biol *25*, 1620-1633.

Kurd, N., and Robey, E.A. (2016). T-cell selection in the thymus: a spatial and temporal perspective. Immunol Rev *271*, 114-126.

Kuwabara, T., Ishikawa, F., Ikeda, M., Ide, T., Kohwi-Shigematsu, T., Tanaka, Y., and Kondo, M. (2021). SATB1-dependent mitochondrial ROS production controls TCR signaling in CD4 T cells. Life Sci Alliance *4*.

Lancaster, J.N., Li, Y., and Ehrlich, L.I.R. (2018). Chemokine-Mediated Choreography of Thymocyte Development and Selection. Trends Immunol *39*, 86-98.

Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfner, K.P., Shirahige, K., and Uhlmann, F. (2006). Establishment of sister chromatid cohesion at the S. cerevisiae replication fork. Mol Cell *23*, 787-799.

Li, L., Leid, M., and Rothenberg, E.V. (2010a). An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. Science *329*, 89-93.

Li, P., Burke, S., Wang, J., Chen, X., Ortiz, M., Lee, S.C., Lu, D., Campos, L., Goulding, D., Ng, B.L.*, et al.* (2010b). Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion. Science *329*, 85-89.

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.

Liu, P., Li, P., and Burke, S. (2010). Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. Immunol Rev *238*, 138-149.

Longabaugh, W.J.R., Zeng, W., Zhang, J.A., Hosokawa, H., Jansen, C.S., Li, L., Romero-Wolf, M., Liu, P., Kueh, H.Y., Mortazavi, A.*, et al.* (2017). Bcl11b and combinatorial resolution of cell fate in the T-cell gene regulatory network. Proc Natl Acad Sci U S A *114*, 5800-5807.

Love, P.E., Lee, J., and Shores, E.W. (2000). Critical relationship between TCR signaling potential and TCR affinity during thymocyte selection. J Immunol *165*, 3080-3087.

Luckey, M.A., Kimura, M.Y., Waickman, A.T., Feigenbaum, L., Singer, A., and Park, J.H. (2014). The transcription factor ThPOK suppresses Runx3 and imposes CD4(+) lineage fate by inducing the SOCS suppressors of cytokine signaling. Nat Immunol *15*, 638-645.

Luckheeram, R.V., Zhou, R., Verma, A.D., and Xia, B. (2012). CD4(+)T cells: differentiation and functions. Clin Dev Immunol *2012*, 925135.

Malhotra, D., Linehan, J.L., Dileepan, T., Lee, Y.J., Purtha, W.E., Lu, J.V., Nelson, R.W., Fife, B.T., Orr, H.T., Anderson, M.S.*, et al.* (2016). Tolerance is established in polyclonal CD4(+) T cells by distinct mechanisms, according to self-peptide expression patterns. Nat Immunol *17*, 187-195.

Mir, R., Pradhan, S.J., Patil, P., Mulherkar, R., and Galande, S. (2016). Wnt/beta-catenin signaling regulated SATB1 promotes colorectal cancer tumorigenesis and progression. Oncogene *35*, 1679-1691.

Miyazaki, M., Miyazaki, K., Chen, K., Jin, Y., Turner, J., Moore, A.J., Saito, R., Yoshida, K., Ogawa, S., Rodewald, H.R.*, et al.* (2017). The E-Id Protein Axis Specifies Adaptive Lymphoid Cell Identity and Suppresses Thymic Innate Lymphoid Cell Development. Immunity *46*, 818-834 e814.

Miyazaki, M., Rivera, R.R., Miyazaki, K., Lin, Y.C., Agata, Y., and Murre, C. (2011). The opposing roles of the transcription factor E2A and its antagonist Id3 that orchestrate and enforce the naive fate of T cells. Nat Immunol *12*, 992-1001.

Molina, T.J., Kishihara, K., Siderovski, D.P., van Ewijk, W., Narendran, A., Timms, E., Wakeham, A., Paige, C.J., Hartmann, K.U., Veillette, A.*, et al.* (1992). Profound block in thymocyte development in mice lacking p56lck. Nature *357*, 161-164.

Murre, C. (2019). Helix-loop-helix proteins and the advent of cellular diversity: 30 years of discovery. Genes Dev *33*, 6-25.

Musio, A. (2020). The multiple facets of the SMC1A gene. Gene *743*, 144612.

Naito, T., Ise, M., Tanaka, Y., Kohwi-Shigematsu, T., and Kondo, M. (2023). Crucial Roles of SATB1 in Regulation of Thymocyte Migration after Positive Selection. J Immunol *211*, 209-218.

Nakayama, T., Hirahara, K., Onodera, A., Endo, Y., Hosokawa, H., Shinoda, K., Tumes, D.J., and Okamoto, Y. (2017). Th2 Cells in Health and Disease. Annu Rev Immunol *35*, 53-84.

Nasmyth, K. (2001). Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. Annu Rev Genet *35*, 673-745.

Nasmyth, K., and Haering, C.H. (2005). The structure and function of SMC and kleisin complexes. Annu Rev Biochem *74*, 595-648.

Neumeister, E.N., Zhu, Y., Richard, S., Terhorst, C., Chan, A.C., and Shaw, A.S. (1995). Binding of ZAP-70 to phosphorylated T-cell receptor zeta and eta enhances its

autophosphorylation and generates specific binding sites for SH2 domain-containing proteins. Mol Cell Biol *15*, 3171-3178.

Nitta, T., Murata, S., Ueno, T., Tanaka, K., and Takahama, Y. (2008). Thymic microenvironments for T-cell repertoire formation. Adv Immunol *99*, 59-94.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Notani, D., Gottimukkala, K.P., Jayani, R.S., Limaye, A.S., Damle, M.V., Mehta, S., Purbey, P.K., Joseph, J., and Galande, S. (2010). Global regulator SATB1 recruits beta-catenin and regulates T(H)2 differentiation in Wnt-dependent manner. PLoS Biol *8*, e1000296.

Nussing, S., Miosge, L.A., Lee, K., Olshansky, M., Barugahare, A., Roots, C.M., Sontani, Y., Day, E.B., Koutsakos, M., Kedzierska, K.*, et al.* (2022). SATB1 ensures appropriate transcriptional programs within naive CD8(+) T cells. Immunol Cell Biol *100*, 636-652.

Pai, S.Y., Truitt, M.L., Ting, C.N., Leiden, J.M., Glimcher, L.H., and Ho, I.C. (2003). Critical roles for transcription factor GATA-3 in thymocyte development. Immunity *19*, 863-875.

Palacios, E.H., and Weiss, A. (2004). Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. Oncogene *23*, 7990-8000.

Park, J.H., Adoro, S., Guinter, T., Erman, B., Alag, A.S., Catalfamo, M., Kimura, M.Y., Cui, Y., Lucas, P.J., Gress, R.E.*, et al.* (2010). Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. Nat Immunol *11*, 257-264.

Patta, I., Madhok, A., Khare, S., Gottimukkala, K.P., Verma, A., Giri, S., Dandewad, V., Seshadri, V., Lal, G., Misra-Sen, J.*, et al.* (2020). Dynamic regulation of chromatin organizer SATB1 via TCR-induced alternative promoter switch during T-cell development. Nucleic Acids Res *48*, 5873-5890.

Pavan Kumar, P., Purbey, P.K., Sinha, C.K., Notani, D., Limaye, A., Jayani, R.S., and Galande, S. (2006). Phosphorylation of SATB1, a global gene regulator, acts as a molecular switch regulating its transcriptional activity in vivo. Mol Cell *22*, 231-243.

Pekowska, A., Klaus, B., Xiang, W., Severino, J., Daigle, N., Klein, F.A., Oles, M., Casellas, R., Ellenberg, J., Steinmetz, L.M.*, et al.* (2018). Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. Cell Syst *7*, 482-495 e410.

Pereira de Sousa, A., Berthault, C., Granato, A., Dias, S., Ramond, C., Kee, B.L., Cumano, A., and Vieira, P. (2012). Inhibitors of DNA binding proteins restrict T cell potential by repressing Notch1 expression in Flt3-negative common lymphoid progenitors. J Immunol *189*, 3822-3830.

Pobezinsky, L.A., Angelov, G.S., Tai, X., Jeurling, S., Van Laethem, F., Feigenbaum, L., Park, J.H., and Singer, A. (2012). Clonal deletion and the fate of autoreactive thymocytes that survive negative selection. Nat Immunol *13*, 569-578.

Purbey, P.K., Singh, S., Kumar, P.P., Mehta, S., Ganesh, K.N., Mitra, D., and Galande, S. (2008). PDZ domain-mediated dimerization and homeodomain-directed specificity are required for high-affinity DNA binding by SATB1. Nucleic Acids Res *36*, 2107-2122.

Purbey, P.K., Singh, S., Notani, D., Kumar, P.P., Limaye, A.S., and Galande, S. (2009). Acetylation-dependent interaction of SATB1 and CtBP1 mediates transcriptional repression by SATB1. Mol Cell Biol *29*, 1321-1337.

Radtke, F., Wilson, A., Stark, G., Bauer, M., van Meerwijk, J., MacDonald, H.R., and Aguet, M. (1999). Deficient T cell fate specification in mice with an induced inactivation of Notch1. Immunity *10*, 547-558.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D.*, et al.* (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320 e324.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C.*, et al.* (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc *14*, 482-517.

Riggs, A.D. (1990). DNA methylation and late replication probably aid cell memory, and type I DNA reeling could aid chromosome folding and enhancer function. Philos Trans R Soc Lond B Biol Sci *326*, 285-297.

Robey, E.A., Fowlkes, B.J., Gordon, J.W., Kioussis, D., von Boehmer, H., Ramsdell, F., and Axel, R. (1991). Thymic selection in CD8 transgenic mice supports an instructive model for commitment to a CD4 or CD8 lineage. Cell *64*, 99-107.

Rothenberg, E.V. (2011). T cell lineage commitment: identity and renunciation. J Immunol *186*, 6649-6655.

Rothenberg, E.V., Moore, J.E., and Yui, M.A. (2008). Launching the T-cell-lineage developmental programme. Nat Rev Immunol *8*, 9-21.

Ruterbusch, M., Pruner, K.B., Shehata, L., and Pepper, M. (2020). In Vivo CD4(+) T Cell Differentiation and Function: Revisiting the Th1/Th2 Paradigm. Annu Rev Immunol *38*, 705-725.

Samelson, L.E. (2002). Signal transduction mediated by the T cell antigen receptor: the role of adapter proteins. Annu Rev Immunol *20*, 371-394.

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J.*, et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A *112*, E6456-6465.

Satoh, Y., Yokota, T., Sudo, T., Kondo, M., Lai, A., Kincade, P.W., Kouro, T., Iida, R., Kokame, K., Miyata, T.*, et al.* (2013). The Satb1 protein directs hematopoietic stem cell differentiation toward lymphoid lineages. Immunity *38*, 1105-1115.

Schebesta, A., McManus, S., Salvagiotto, G., Delogu, A., Busslinger, G.A., and Busslinger, M. (2007). Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. Immunity *27*, 49-63.

Schmitt, T.M., and Zuniga-Pflucker, J.C. (2002). Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. Immunity *17*, 749-756.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Scripture-Adams, D.D., Damle, S.S., Li, L., Elihu, K.J., Qin, S., Arias, A.M., Butler, R.R., 3rd, Champhekar, A., Zhang, J.A., and Rothenberg, E.V. (2014). GATA-3 dose-dependent checkpoints in early T cell commitment. J Immunol *193*, 3470-3491.

Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G.*, et al.* (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res *23*, 2066-2077.

Seitan, V.C., Hao, B., Tachibana-Konwalski, K., Lavagnolli, T., Mira-Bontenbal, H., Brown, K.E., Teng, G., Carroll, T., Terry, A., Horan, K.*, et al.* (2011). A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature *476*, 467-471.

Seitan, V.C., Krangel, M.S., and Merkenschlager, M. (2012). Cohesin, CTCF and lymphocyte antigen receptor locus rearrangement. Trends Immunol *33*, 153-159.

Seo, W., and Taniuchi, I. (2016). Transcriptional regulation of early T-cell development in the thymus. Eur J Immunol *46*, 531-538.

Setoguchi, R., Tachibana, M., Naoe, Y., Muroi, S., Akiyama, K., Tezuka, C., Okuda, T., and Taniuchi, I. (2008). Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. Science *319*, 822-825.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shah, K., Al-Haidari, A., Sun, J., and Kazi, J.U. (2021). T cell receptor (TCR) signaling in health and disease. Signal Transduct Target Ther *6*, 412.

Shinzawa, M., Moseman, E.A., Gossa, S., Mano, Y., Bhattacharya, A., Guinter, T., Alag, A., Chen, X., Cam, M., McGavern, D.B.*, et al.* (2022). Reversal of the T cell immune system reveals the molecular basis for T cell lineage fate determination in the thymus. Nat Immunol *23*, 731-742.

Sidwell, T., and Rothenberg, E.V. (2021). Epigenetic Dynamics in the Function of T-Lineage Regulatory Factor Bcl11b. Front Immunol *12*, 669498.

Singer, A. (2002). New perspectives on a developmental dilemma: the kinetic signaling model and the importance of signal duration for the CD4/CD8 lineage decision. Curr Opin Immunol *14*, 207-215.

Singer, A., Adoro, S., and Park, J.H. (2008). Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. Nat Rev Immunol *8*, 788-801.

Singh, H., Khan, A.A., and Dinner, A.R. (2014). Gene regulatory networks in the immune system. Trends Immunol *35*, 211-218.

Smith-Garvin, J.E., Koretzky, G.A., and Jordan, M.S. (2009). T cell activation. Annu Rev Immunol *27*, 591-619.

Stadhouders, R., Filion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. Nature *569*, 345-354.

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y*., et al.* (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat Genet *50*, 238-249.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003). Positive and negative selection of T cells. Annu Rev Immunol *21*, 139-176.

Stritesky, G.L., Jameson, S.C., and Hogquist, K.A. (2012). Selection of self-reactive T cells in the thymus. Annu Rev Immunol *30*, 95-114.

Strom, L., Karlsson, C., Lindroos, H.B., Wedahl, S., Katou, Y., Shirahige, K., and Sjogren, C. (2007). Postreplicative formation of cohesion is required for repair and induced by a single DNA break. Science *317*, 242-245.

Sun, G., Liu, X., Mercado, P., Jenkinson, S.R., Kypriotou, M., Feigenbaum, L., Galera, P., and Bosselut, R. (2005). The zinc finger protein cKrox directs CD4 lineage differentiation during intrathymic T cell positive selection. Nat Immunol *6*, 373-381.

Sun, L., Su, Y., Jiao, A., Wang, X., and Zhang, B. (2023). T cells in health and disease. Signal Transduct Target Ther *8*, 235.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. Sci Adv *5*, eaaw1668.

Takahama, Y. (2006). Journey through the thymus: stromal guides for T-cell development and selection. Nat Rev Immunol *6*, 127-135.

Taniuchi, I., Osato, M., Egawa, T., Sunshine, M.J., Bae, S.C., Komori, T., Ito, Y., and Littman, D.R. (2002). Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. Cell *111*, 621-633.

Tao, X., Constant, S., Jorritsma, P., and Bottomly, K. (1997). Strength of TCR signal determines the costimulatory requirements for Th1 and Th2 CD4+ T cell differentiation. J Immunol *159*, 5956-5963.

Thompson, P.K., and Zuniga-Pflucker, J.C. (2011). On becoming a T cell, a convergence of factors kick it up a Notch along the way. Semin Immunol *23*, 350-359.

Tiemessen, M.M., Baert, M.R., Schonewille, T., Brugman, M.H., Famili, F., Salvatori, D.C., Meijerink, J.P., Ozbek, U., Clevers, H., van Dongen, J.J*., et al.* (2012). The nuclear effector of Wnt-signaling, Tcf1, functions as a T-cell-specific tumor suppressor for development of lymphomas. PLoS Biol *10*, e1001430.

van Oers, N.S., Killeen, N., and Weiss, A. (1996). Lck regulates the tyrosine phosphorylation of the T cell receptor subunits and ZAP-70 in murine thymocytes. J Exp Med *183*, 1053-1062.

Wang, H., Kadlecek, T.A., Au-Yeung, B.B., Goodfellow, H.E., Hsu, L.Y., Freedman, T.S., and Weiss, A. (2010). ZAP-70: an essential kinase in T-cell signaling. Cold Spring Harb Perspect Biol *2*, a002279.

Wang, L., Wildt, K.F., Zhu, J., Zhang, X., Feigenbaum, L., Tessarollo, L., Paul, W.E., Fowlkes, B.J., and Bosselut, R. (2008). Distinct functions for the transcription factors GATA-3 and ThPOK during intrathymic differentiation of CD4(+) T cells. Nat Immunol *9*, 1122-1130.

Wang, Z., Yang, X., Chu, X., Zhang, J., Zhou, H., Shen, Y., and Long, J. (2012). The structural basis for the oligomerization of the N-terminal domain of SATB1. Nucleic Acids Res *40*, 4193-4202.

Weber, B.N., Chi, A.W., Chavez, A., Yashiro-Ohtani, Y., Yang, Q., Shestova, O., and Bhandoola, A. (2011). A critical role for TCF-1 in T-lineage specification and differentiation. Nature *476*, 63-68.

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 e1528.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Will, B., Vogler, T.O., Bartholdy, B., Garrett-Bakelman, F., Mayer, J., Barreyro, L., Pandolfi, A., Todorova, T.I., Okoye-Okafor, U.C., Stanley, R.F.*, et al.* (2013). Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. Nat Immunol *14*, 437-445.

Wojciechowski, J., Lai, A., Kondo, M., and Zhuang, Y. (2007). E2A and HEB are required to block thymocyte proliferation prior to pre-TCR expression. J Immunol *178*, 5717-5726.

Wolfer, A., Wilson, A., Nemir, M., MacDonald, H.R., and Radtke, F. (2002). Inactivation of Notch1 impairs VDJbeta rearrangement and allows pre-TCR-independent survival of early alpha beta Lineage Thymocytes. Immunity *16*, 869-879.

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J.*, et al.* (2017). Topologically associating

domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J *36*, 3573-3599.

Yang, Q., Jeremiah Bell, J., and Bhandoola, A. (2010). T-cell lineage determination. Immunol Rev *238*, 12-22.

Yang, Q., Kardava, L., St Leger, A., Martincic, K., Varnum-Finney, B., Bernstein, I.D., Milcarek, C., and Borghesi, L. (2008). E47 controls the developmental integrity and cell cycle quiescence of multipotential hematopoietic progenitors. J Immunol *181*, 5885-5894.

Yashiro-Ohtani, Y., He, Y., Ohtani, T., Jones, M.E., Shestova, O., Xu, L., Fang, T.C., Chiang, M.Y., Intlekofer, A.M., Blacklow, S.C*., et al.* (2009). Pre-TCR signaling inactivates Notch1 transcription by antagonizing E2A. Genes Dev *23*, 1665-1676.

Yashiro-Ohtani, Y., Ohtani, T., and Pear, W.S. (2010). Notch regulation of early thymocyte development. Semin Immunol *22*, 261-269.

Yassai, M., Ammon, K., Goverman, J., Marrack, P., Naumov, Y., and Gorski, J. (2002). A molecular marker for thymocyte-positive selection: selection of CD4 single-positive thymocytes with shorter TCRB CDR3 during T cell development. J Immunol *168*, 3801-3807.

Yasuda, K., Kitagawa, Y., Kawakami, R., Isaka, Y., Watanabe, H., Kondoh, G., Kohwi-Shigematsu, T., Sakaguchi, S., and Hirota, K. (2019). Satb1 regulates the effector program of encephalitogenic tissue Th17 cells in chronic inflammation. Nat Commun *10*, 549.

Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P., and Kohwi-Shigematsu, T. (2002). SATB1 targets chromatin remodelling to regulate genes over long distances. Nature *419*, 641-645.

Yoshida, H., Lareau, C.A., Ramirez, R.N., Rose, S.A., Maier, B., Wroblewska, A., Desland, F., Chudnovskiy, A., Mortha, A., Dominguez, C*., et al.* (2019). The cis-Regulatory Atlas of the Mouse Immune System. Cell *176*, 897-912 e820.

Yui, M.A., and Rothenberg, E.V. (2014). Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol *14*, 529-545.

Zeidan, N., Damen, H., Roy, D.C., and Dave, V.P. (2019). Critical Role for TCR Signal Strength and MHC Specificity in ThPOK-Induced CD4 Helper Lineage Choice. J Immunol *202*, 3211-3225.

Zelenka, T., Klonizakis, A., Tsoukatou, D., Papamatheakis, D.A., Franzenburg, S., Tzerpos, P., Tzonevrakis, I.R., Papadogkonas, G., Kapsetaki, M., Nikolaou, C*., et al.* (2022). The 3D enhancer network of the developing T cell genome is shaped by SATB1. Nat Commun *13*, 6954.

Zerrahn, J., Held, W., and Raulet, D.H. (1997). The MHC reactivity of the T cell repertoire prior to positive and negative selection. Cell *88*, 627-636.

Zhang, J.A., Mortazavi, A., Williams, B.A., Wold, B.J., and Rothenberg, E.V. (2012). Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. Cell *149*, 467-482.

Zhang, S.L., and Bhandoola, A. (2014). Trafficking to the thymus. Curr Top Microbiol Immunol *373*, 87-111.

Zheng, H., and Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol *20*, 535-550.

Zhou, L., Chong, M.M., and Littman, D.R. (2009). Plasticity of CD4+ T cell lineage differentiation. Immunity *30*, 646-655.

Zook, E.C., Li, Z.Y., Xu, Y., de Pooter, R.F., Verykokakis, M., Beaulieu, A., Lasorella, A., Maienschein-Cline, M., Sun, J.C., Sigvardsson, M.*, et al.* (2018). Transcription factor ID2 prevents E proteins from enforcing a naive T lymphocyte gene program during NK cell development. Sci Immunol *3*.

Chapter 2. Identification and characterization of liquid-like condensate formation by Satb1 in thymocytes

2.1 Introduction

In a highly packaged genome, activation of a subset of genes in a cell type specific manner is a cumbersome task and requires a concerted effort by different protein-rich machineries helped by regulatory genomic elements. Transcription factors typically cooperate with other proteins on the chromatin to regulate transcription, by establishing or maintaining contacts of Pol II occupied promoters with proximal/distal enhancers. TFs and other regulatory proteins interact physically with each other to surround gene regulatory elements such as enhancers as well as promoters in the nucleus, forming a "hub," or region of higher local protein concentration. (Hnisz et al., 2017). Formation of transcriptional ‘hubs’ is suggested to more efficiently promote the process of transcription. Recent evidences have demonstrated that the interactions forming these locally increased protein concentrations are not one-to-one. Weak protein-protein associations may be the mediating factor for this more intricate kind of cooperation, referred to as non-stoichiometric interaction. Because these interactions frequently involve proteins that are intrinsically disordered or contains low-complexity regions (Chong et al., 2018). The dynamics of transcription initiation as well as the recently discovered involvement of phase separation in transcription as well as genome packaging is discussed below.

2.1.1 Transcriptional activation

RNA polymerase II (Pol II), an enzyme that transcribes eukaryotic protein-coding mRNAs and untranslated RNA types, require a precise spatio-temporal control which involves an interplay of a number of factors and co-activators ((Roeder, 1998). Among them are a group of general transcription factors (GTFs) that Pol II needs in order to initiate transcription initiation specific to a promoter (Schier and Taatjes, 2020; Thomas and Chiang, 2006). A crucial rate-limiting event in gene transcription is the assembly and function of the pre-initiation complex (PIC), which is formed at the promoter's core by Pol II and these GTFs (Roeder, 1998). Most commonly used promoter element is the ‘TATA’ sequence. TATA binding protein (TBP) binding to the TATA site initiates pre-initiation complex (PIC) assembly on a promoter containing TATA elements (Buratowski et al., 1988; Kao et al., 1990), which becomes a docking site for other GTFs. Co-activators such as TFIID and the mediator complex interact with the PIC, and help recruitment of cis-regulatory elements such as enhancers, driving cell type specific gene regulation (Figure 2.1.1) (Black et al., 2006; Richter et al., 2022).

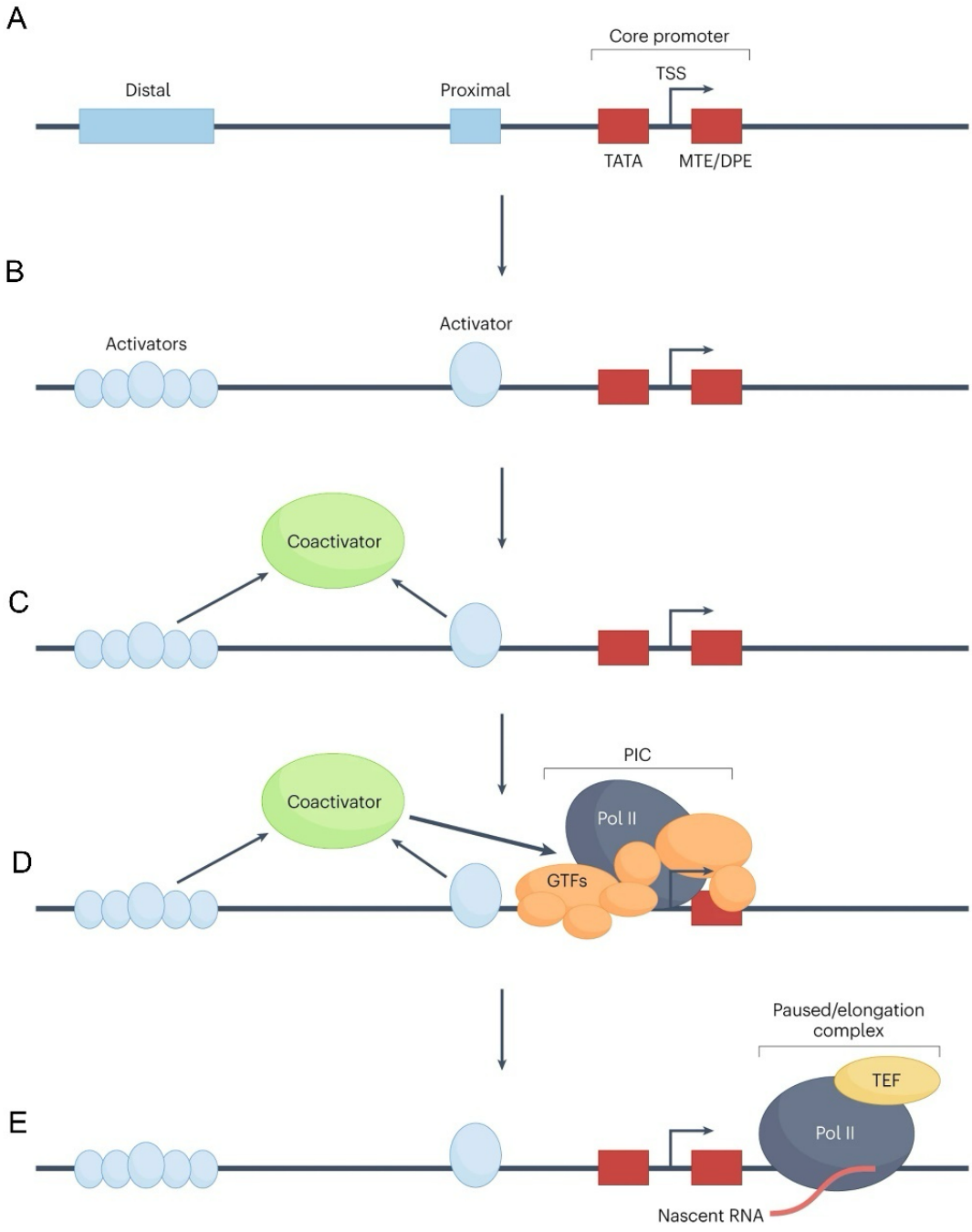


Figure 2.1.1: Pre-initiation complex (PIC) formation involves multiple steps. A. Shows a cognate template region that includes a variety of regulatory sequence elements that are not usually present in a single gene locus. **B.** Signal-induced transcription factors may bind to sites close to the core-promoter motifs or sites found in distal enhancers, such as super-enhancers (SE). C. The transcription start site (TSS) is specified by the combination of core promoter elements. They could also contain downstream promoter elements (DPEs) in addition to the TATA box sequence. **D.** A series of actions, including the recruitment of one or

more factors, are triggered by the recruitment of activator protein. **E.** PIC promotes nascent RNA transcription and elongation by Pol II. Reproduced from Malik and Roeder, 2023.

2.1.2 Chromatin organization and gene regulation

In a genome with a tightly packed chromatin, regulatory complexes that regulate RNA transcription must dynamically rearrange the chromatin in order to reach promoter elements (Woodcock and Ghosh, 2010). Accurate execution of such genome-wide regulation mechanisms requires a temporally controlled genome re-organization. Important aspects of genome organization include local segregation of the chromatin into architectural domains determined by DNA and histone modifications, as well as topological modulation of the chromatin via molecular machineries in an ATP-dependent manner (Bonev and Cavalli, 2016; Shaban et al., 2018). A few studies indicate that genome organization differ across an organism's cell types, pointing to an unchanging aspect of genome structure (de Wit et al., 2013; Nora et al., 2012). On the other hand, the chromatin associations among the various regulatory elements appear to be unique to a cell type, enabling for the control of transcription tailored to a particular cell type (Weintraub et al., 2017). These interactions occur within the bigger contact domain (as discussed below) and they consist of distant enhancer-promoter interactions (Javierre et al., 2016; Stadhouders et al., 2018). Given that a number of disease-associated mutations happen close to the enhancers, such cell type-intrinsic enhancer-promoter contacts are crucial. In developing T cells, critical interactions within the ‘active’ chromatin is mediated by looping factors such as SATB1 in a signal dependent manner (Galande et al., 2007).

2.1.3 Chromosomal compartments and topologically associated domains (TADs)

To accurately allow access of a ‘future’ transcriptionally active gene promoter and to mediate interactions between various regulatory elements, a hierarchical organization of the chromatin is established. The chromatin is compressed throughout intermediate genome length scales (of the range 10 Kb–1 Mb) and needs to constantly probe its immediate environment in order for two distal genomic regions to interact during a functionally meaningful period (Di Pierro et al., 2018). Each chromosome roughly occupies a distinct volume in an interphase nucleus as shown by 3D chromosome painting using fluorescent in situ hybridization (3D FISH). These are described as chromosomal territories (CTs) (Cremer et al., 2008). At the mega-base scale of sub-chromosomal level in the eukaryotic nuclei, two physically distinct and functionally segregated compartments occur: A compartment, which is functionally active and is generally

gene-rich and harbours active histone modification H3K36me3, and an inactive B compartment, which is usually gene-poor and contains a repressive H3K9me3 mark (Lieberman-Aiden et al., 2009; Zhang et al., 2012) (Figure 2.1.2).

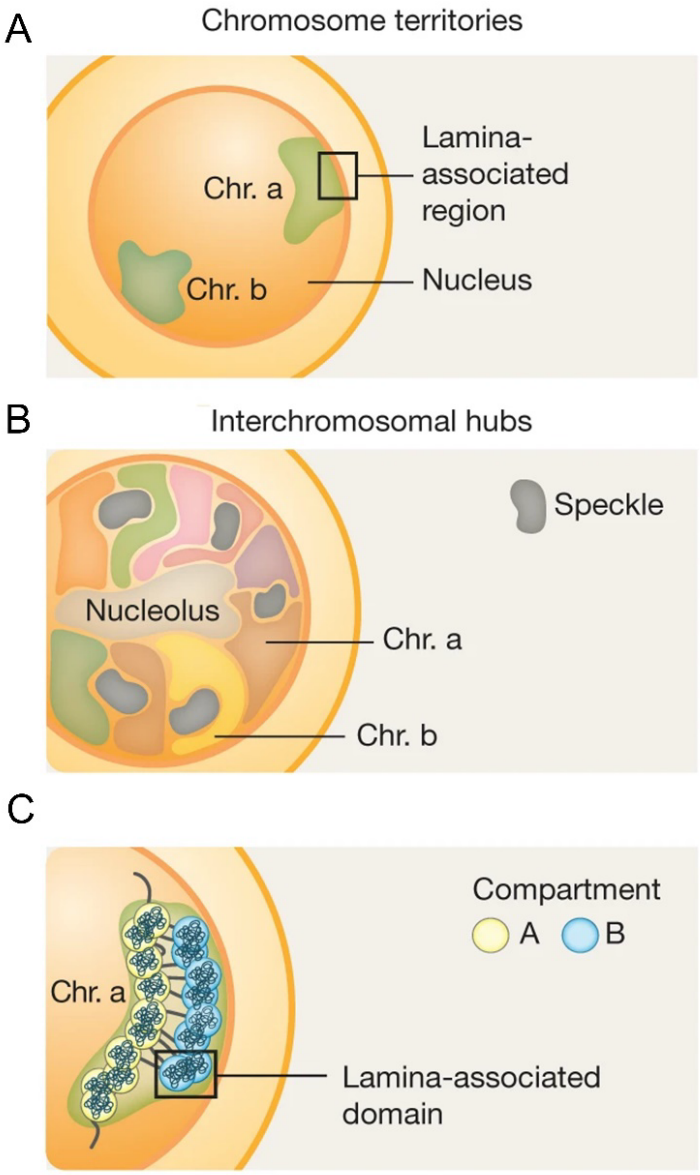


Figure 2.1.2: Distinct principles that govern chromatin packaging within the nucleus. A. Different chromosomes (labelled as Chr a and Chr b) comprise different regions inside the nucleus. **B.** These chromosomes form the interchromosomal hubs, wherein the inactive chromatin is closest to the nucleolus and the active chromatin is located nearby nuclear speckles. **C.** Active and inactive chromatin partition into A and B compartments inside the chromosomal territories, respectively. Reproduced from Stadhouders et al., 2019.

Further, megabase-scale segments of DNA sequences known as topologically associated domains (TADs) or contact domain have been identified using Hi-C contact matrices. The interactions among the sequences inside the segment had significantly greater frequencies than those between the sequences outside of it. Boundaries of these blocks are often richer in the SMC complex cohesin and CCCTC-binding factor (CTCF) define a TAD (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). These self-interacting TADs are formed via ‘loop-extrusion’ through SMC complexes with the help of CTCF which acts like a barrier to the extrusion, thus controlling the loop length (Fudenberg et al., 2016; Sanborn et al., 2015). Loop extrusion could be diminished by depleting cohesin (core subunits or its 'loading' proteins), which results in strengthened intrachromosomal compartmentalization along with the loss of TAD structure throughout the genome. This suggests that loop extrusion and compartmentalization utilize distinct mechanisms that may operate antagonistically. Therefore, it seems plausible that compartmentalization may be the primary strategy of three-dimensional chromatin organization, while loop extrusion creates isolated genomic segments which are impenetrable to subsequent compartmentalization (Figure 2.1.3) (Nuebler et al., 2018; Schwarzer et al., 2017). TADs generally guide the interactions locally among gene regulatory elements via restricting contacts of promoters and enhancers mostly at the inside of particular TADs (Dowen et al., 2014; Shen et al., 2012; Sun et al., 2019).

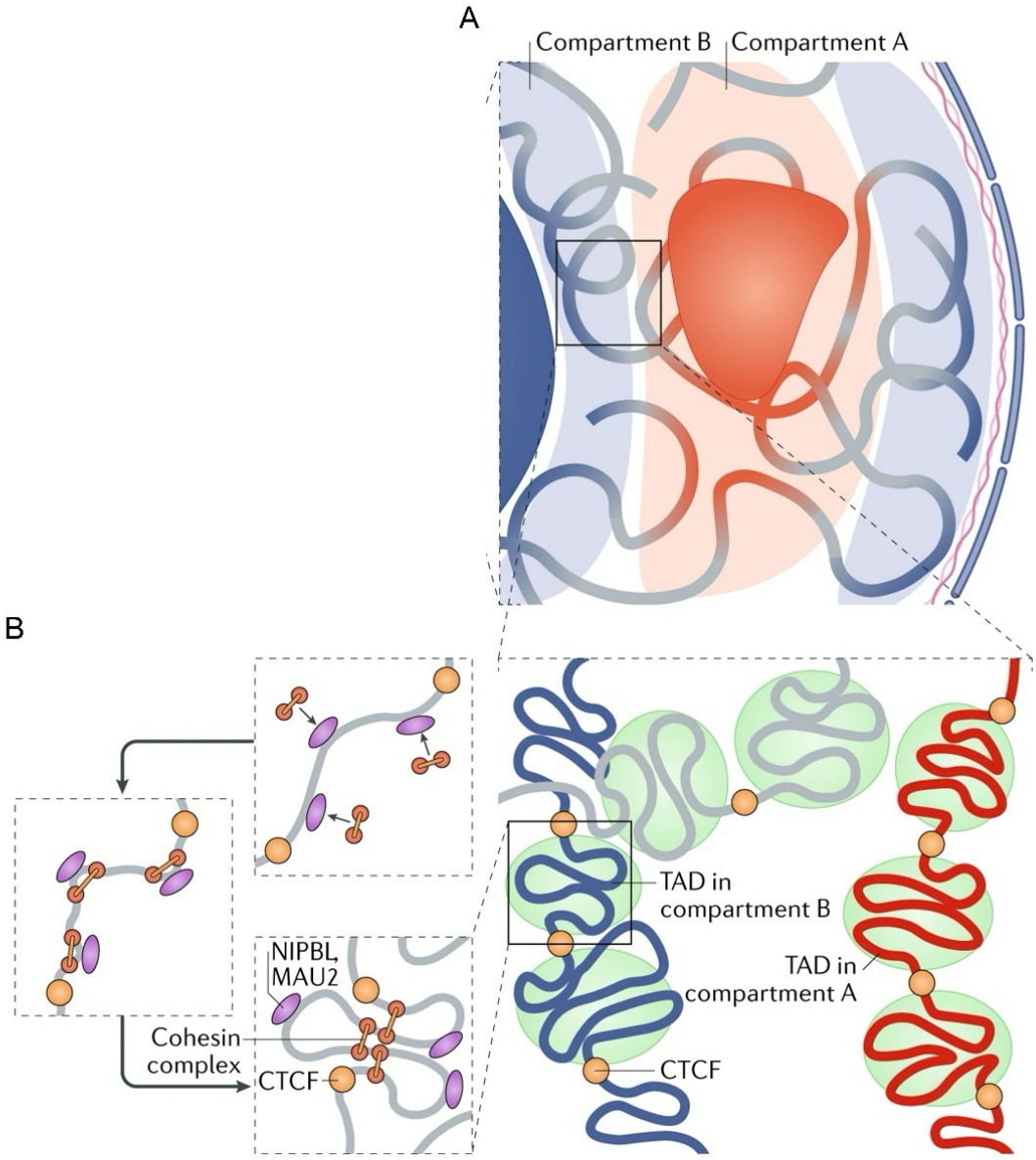


Figure 2.1.3: Formation of topologically associated domains (TADs). A. On a smaller level, transcriptionally engaged areas interact with other active regions to create compartment A. Inactive areas combine with other inactive regions to form compartment B. These inactive regions are predominantly linked with the nuclear lamina as well as to the nucleolus. **B.** Topologically associating domains (TADs) are formed locally when genomic areas exhibit strong self-interactions and are isolated from neighboring genomic regions. TAD borders are the preferred location of CCCTC-binding factor (CTCF) binding on the chromatin. Chromatin looping is facilitated inside each TAD by loop domains that are controlled by the cohesin complex, which extrudes DNA through its ring-like conformation. It is loaded onto the DNA by nipped-B-like protein (NIPBL). CTCF can obstruct cohesive sliding, leading to loop boundary formation. Reproduced from Zheng and Xie, 2019.

2.1.4 Impact of genome conformation on cell fate decisions

As chromatin self-organizes as non-stochastic 3D globules -TADs, it implies that a functional outcome in terms of gene regulation and cell fate decision making is probable (Figure 2.1.4). Cell type influences the configuration of the genome, at least in part. For example, different cell types have varying position of chromosomal territories, even at the areas of mixing (Fraser et al., 2015). It was shown that in the course of cell differentiation as well as during reprogramming, as much as 35% of the chromatin flips between the A and B compartments (Bonev et al., 2017; Pekowska et al., 2018; Rao et al., 2014; Stadhouders et al., 2018). The most changeable measure of genome conformation, although is the promoter-enhancer contacts. It has been found that there is a substantial level of cell-state distinctiveness (~80%) in the probability of particular promoter-enhancer connections to occur (Bonev and Cavalli, 2016; Javierre et al., 2016; Kieffer-Kwon et al., 2013; Zhang et al., 2013).

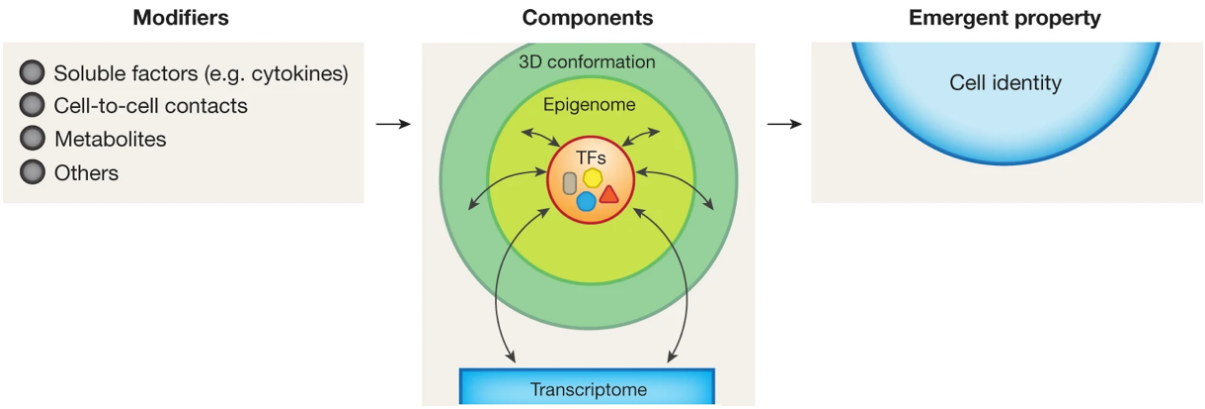


Figure 2.1.4: Schematic depicting induction of cell-fate decisions. The activity of transcription factors (TFs) is modulated by signal transduction triggered by extrinsic factors/states. The chromatin landscape takes on a particular three-dimensional conformation, interacts with the signal induced transcription factors or modifiers. Transcription factors modify the activation program of the cell by enlisting the help of epigenetic regulators and the transcriptional apparatus. Reproduced from Stadhouders et al., 2019.

Furthermore, genome architecture is implicated in transcriptional control directly by the pattern of 3D genome folding during cellular differentiation (Dixon et al., 2015; Hu et al., 2018) (Figure 2.1.5). For instance, it was recently demonstrated that TCF1 promotes T cell progenitor development by modulating the chromatin architecture at key lineage specifying genes (Wang

et al., 2022). Similarly, the formation and preservation of the global chromatin architecture of B cells depend on Pax5 for proper lineage specification (Johanson et al., 2018).

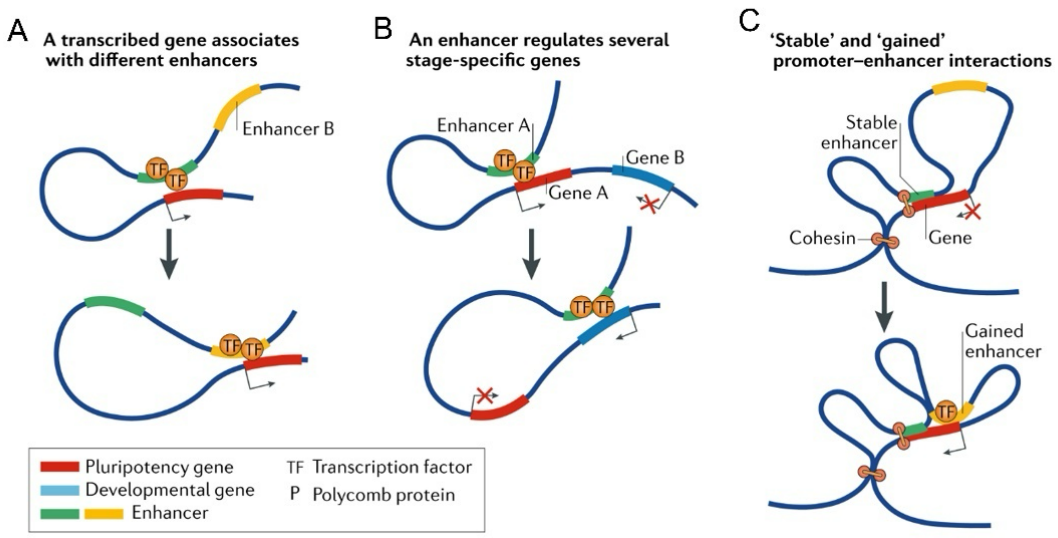


Figure 2.1.5: Interactions among the cis-regulatory elements during development. A. In various cell types, transcriptionally active genes can bind to distinct arrays of enhancers via looping through TFs. **B.** Through interactions between stage-specific promoters and enhancers, different genes that are particular to a stage can be induced by the same enhancers. **C.** In the case of stage-specific transcriptional activation, there are two kinds of promoter–enhancer interactions: Cohesin primarily mediates the stable associations between promoter and enhancer that precede the activation of the target genes, while promoter-enhancer communication that arise upon gene induction are usually unaffected by cohesin. Reproduced from Zheng and Xie, 2019.

2.1.5 Phase separation: An emerging mechanism for genome organization

Extensive transcriptional dysregulation does not occur with the notable impairment of topological insulation upon CTCF or cohesin removal. This implies that the significance of topological domain organization in limiting enhancer activity is relatively restricted (Haarhuis et al., 2017; Nora et al., 2017; Schwarzer et al., 2017). Nonetheless, it has been suggested that clustering regions with similar activity near certain nuclear regions may increase the local concentration of pertinent related components, thereby improving the effectiveness of gene regulation mechanisms (Wang et al., 2016). It has recently been suggested that phase

separation plays a significant part in the organization of the 3D genome (Hnisz et al., 2017; Maeshima et al., 2016).

Condensates, or liquid-like droplets composed of self-organizing proteins, are membrane-free organelles that increase local concentration of some molecules while excluding others, describes the phenomenon of phase separation (Figure 2.1.6) (Hyman et al., 2014; Shin et al., 2018). From a historical standpoint, nucleoli and Cajal bodies have been described as membrane-less organelles (Handwerger et al., 2005; Monneron and Bernhard, 1969). Perinuclear RNA granules or P-bodies were the first membrane-less organelles to be thoroughly studied; they were demonstrated to possess liquid-like characteristics (forming condensates) (Brangwynne et al., 2009). It has only recently been realized that these condensates are common occurrence in biology, and could participate in fine tuning genomic interactions (Mitrea and Kriwacki, 2016).

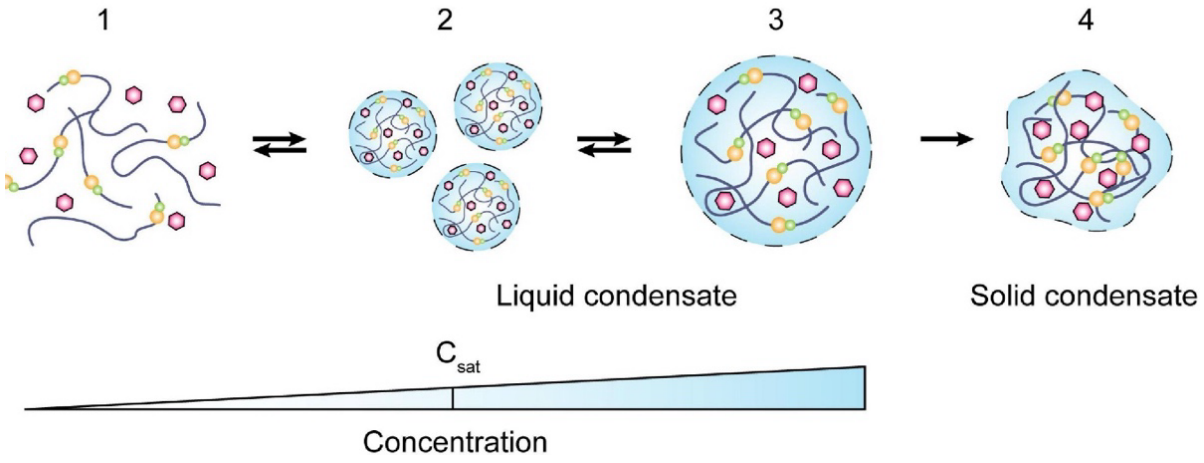


Figure 2.1.6: Schematic representing Liquid-liquid phase separation (LLPS). Various proteins can undergo demixing in an aqueous solution. Demixed phases frequently split into smaller droplets or merge into larger droplets in a generally reversible manner. Depending on the concentration of the protein, the droplets might become solid or gel-like condensates. Reproduced from Gao et al., 2021.

2.1.5.1 Biophysical properties of condensates

Mixtures combining protein-protein, protein-DNA or protein-RNA could constitute biomolecular condensates; every combination has unique biophysical characteristics based on their concentration, affinity, and composition (Feric et al., 2016). Energetically, to raise the system's entropy, a two-component solution typically mixes homogeneously. Phase separation will take place if the interactions involving similar molecules are energetically advantageous enough that demixing lowers the overall free-energy (Figure 2.1.7A). Weak multivalent interactions among intrinsically disordered regions (IDRs) with low complexity amino acid composition allows for phase separation (Wang et al., 2018). Purified IDR containing proteins or IDR domains demix or go through phase transition in a concentration and temperature-dependent manner, in vitro (Figure 2.1.7B). These droplets display characteristics of liquids, such surface tension and nucleation (Kato et al., 2012). Certain condensates can alternate between solid, liquid, and gel-like phases, displaying a metastable dynamic nature (Figure 2.1.7C) based on physical variables like temperature, pH, concentration, and intensity of contact (Dumetz et al., 2008). These condensates exhibit variable scaling, indicating that as droplets develop, dissolve, agglomerate, and age, their size and biophysical characteristics alter (Berry et al., 2018).

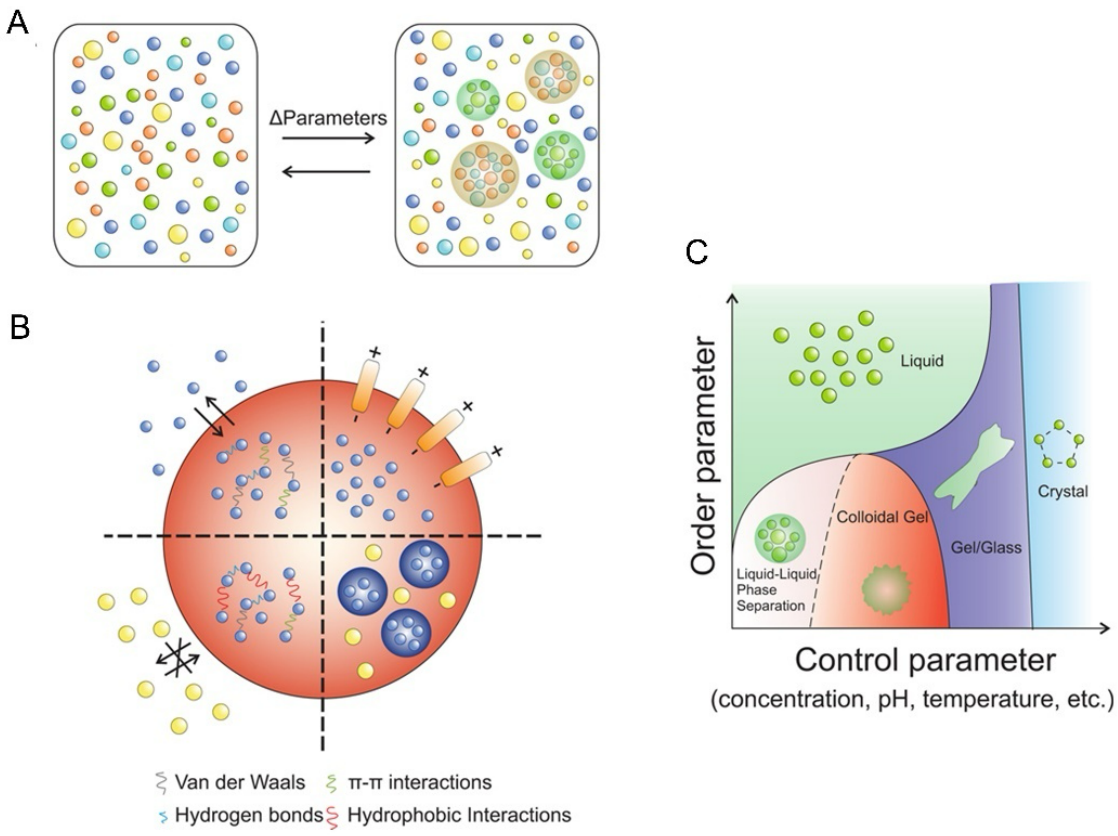


Figure 2.1.7: Modes of phase separation in vitro. A. Numerous regions with greater local protein concentrations as a result 'demixing'. These condensates can be either made up of only one protein type which prevents other species from entering, or they can be made up of several protein types which results in mixed condensates. **B.** Condensates possess unique biophysical characteristics and are created by weak protein-protein interactions which are non-covalent. **C.** phase diagram which is used to identify phase transitions is shown. Biomolecular condensates can exhibit a wide range of biophysical features, resulting in liquid-like droplets, gels, and other structures, depending on various external parameters. Reproduced from Wagh et al., 2021.

2.1.5.2 Phase separation modulates transcription

Transcription is a dynamic process involving the assembly of RNA Pol II machinery, multiple transcription factors and various regulatory elements to specific genomic regions in order to activate gene activation. Apart from the presence of a DNA binding domain (DBD) and the transactivation domain (AD), a large number of transcription factors also harbour intrinsically disordered regions (IDRs), enabling weak multivalent interactions with the other TFs. These IDRs may be divided into families that communicate with one another but keep others apart, allowing specificity to TF condensates (Chong et al., 2018). DBD enables TFs to recruitment to certain enhancers or promoter regions (Figure 2.1.8A), permitting an IDR-mediated nucleation of condensates at particular genomic loci (Wei et al., 2020). TF condensates are also involved in recruiting distal enhancers, forming a super-enhancer hub (Nair et al., 2019; Shin et al., 2018). IDR regions found in the Pol II C-terminal domain (CTD) lead Pol II to phase transition into transcriptional condensates mediated by TF binding and Mediator complex (Figure 2.1.8B and C) (Boehning et al., 2018).

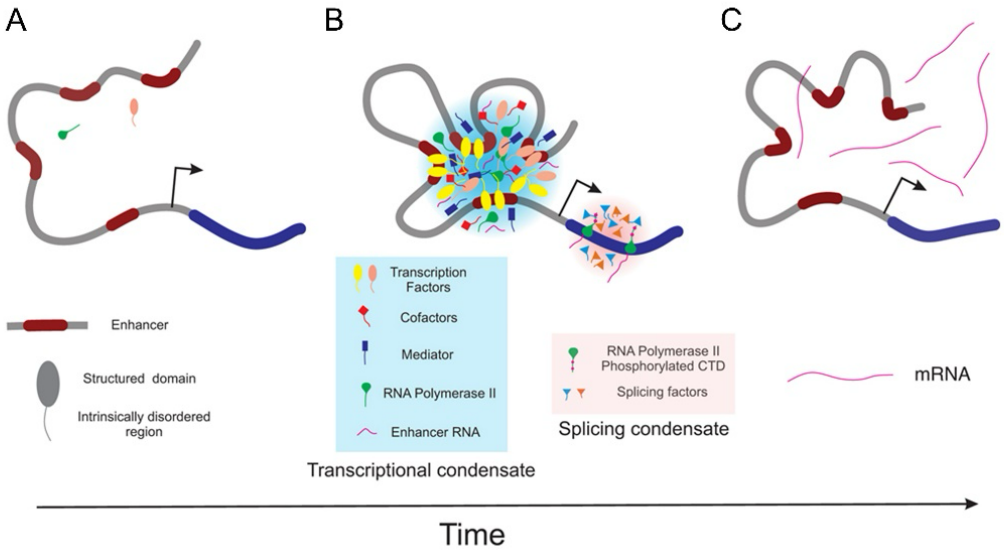


Figure 2.1.8: Phase separation affects transcription dynamics. A. Enhancers and promoter elements are searched by transcription factors for their corresponding DNA motifs. **B.** TF binding results in the creation of transcriptional hubs or condensates. TFs along with RNA Pol II, and nascent mRNA could be present in these condensates **C.** Increased elongated RNA concentration promotes condensate dissolution. Phase behaviour of Pol II is further modulated via phosphorylation of its C terminal domain (CTD). Reproduced from Wagh et al., 2021.

2.1.5.3 Long-range chromatin interactions by phase separation

Varying spatial nuclear structure is seen in cells of the same tissue, demonstrating the tremendous heterogeneity of the 3D genome (Finn et al., 2019), which makes it difficult for TFs to find their cognate binding sites. TFs can be divided into two categories based on their mobility as analyzed using single molecular tracking; chromatin-associated and IDR-mediated restriction states that are in coherence with liquid-liquid phase separation (LLPS) (Garcia et al., 2021). Restricting spontaneous genomic interactions is one of the proposed features of phase separation (Figure 2.1.9). For example, target genes can be spatially clustered directly by NANOG to create a three-dimensional hub (de Wit et al., 2013). Similarly, it was shown that the clusters of repressed or inactive genes have hubs of polycomb, which could possibly be regulated via phase separation. Regulatory proteins may interact with each other using their

IDRs or via homotypic dimerization to generate these three-dimensional hubs (Boija et al., 2018; Chong et al., 2018).

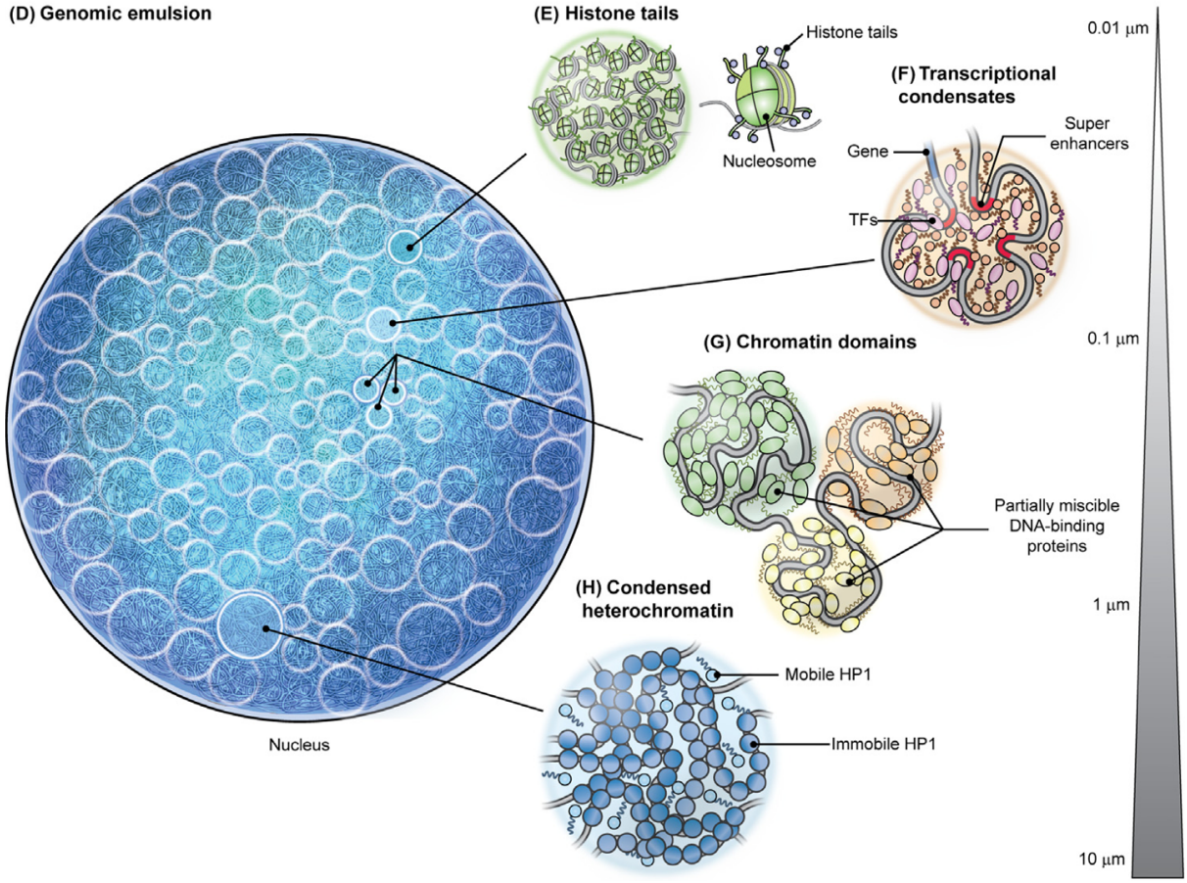


Figure 2.1.9: Phase separation modulates chromatin organization at different levels. The associations among histone tails, as well as transcriptional condensates comprising enhancers and TFs, are examples of chromatin organization control at different length scales. Reproduced from Feric and Misteli, 2021.

In this chapter, we describe the property of Satb1 to undergo condensate formation. We first identify the degree of disorder present in the amino acid makeup of Satb1. We then assessed the biophysical properties of such as concentration, nucleic acid binding and liquidity in vitro. Furthermore, we also probed into the specific domains that are involved in demixing behaviour of Satb1. In multiple stages of thymocyte development, we show that Satb1 forms nuclear foci,

which are concordant with phase behaviour. We also tested the dynamic nature of these foci in DP thymocytes. We demonstrate that Satb1 nuclear foci co-localize with Smc1a, which also display a foci-like pattern. Whereas Ctcf show a more diffused localization in various stages of T cell development. We finally show that specific amino acid residues of Satb1 implicated in neurological impairments and cancer are important for driving the phase transition of Satb1 in vitro. Collectively, Satb1 mediated regulation of transcription and genome structure may be promoted by its ability to undergo condensate formation.

2.2 Materials and Methods

2.2.1 Mice

Six- to ten-week-old WT Satb1 fl/fl, CD4cre Satb1 fl/fl and VaviCreERT2 Satb1 fl/fl mice were used to prepare the single cell suspension of thymocytes for the flow cytometry analysis as well as for sorting the subpopulations of thymocytes which are in different stages of T-cell development. All the strains were bred and maintained in sterile environment and the dissections were carried out in accordance with the guidelines of NFGFHD, IISER Pune.

2.2.2 Cell sorting

The suspension of thymocytes was prepared using thymii from 6-week-old mice from various genotypes indicated and were used for surface staining. Before staining, Fc -blocking was done with anti-CD16.32 antibody (BD Biosciences). Then, the thymocytes were subjected to surface staining using the following the fluorochrome tagged antibodies: BV 421 anti-mouse CD4 (Biolegend); PE anti-mouse CD8a (Clone 53-6.7, BD Biosciences). The thymocyte sub populations such as CD4-CD8- double negative (DN), $CD4^+CD8^+$ double positive (DP), $CD4^+$ SP thymocytes (total CD4SP), and CD8+ SP thymocytes (total CD8SP) were FACS sorted using FACS Aria III SORP (BD biosciences).

2.2.3 Immunostaining

A single-cell suspension was obtained from thymi of WT and Satb1 cKO mice that were six weeks old. In order to fix thymocytes, 2% paraformaldehyde was used. Permeabilization of fixed thymocytes was performed with 0.1% Triton X-100. Following permeabilization, thymocytes were treated for three hours at room temperature with mouse anti-SATB1 (Abcam), goat anti-CTCF (CST) and rabbit anti-SMC1A (Abcam). Thymocytes were stained intracellularly and then washed with 1X PBS containing 0.01% Tween 20. The secondary antibodies were stained for one hour at room temperature using the fluorochrome-tagged secondary antibodies, anti-mouse GFP 488, anti-goat Alexafluor (AF) 568 and anti-rabbit AF 594. Following the staining with a secondary antibody, DNA was labeled with DAPI (Sigma). Individual images and Z-stack pictures were obtained and cells were seen using a SP8 microscope (Leica Biosystems).

2.2.4 Protein expression and purification

The expression plasmid pTRIEX 3 neo containing the SATB1 full-length coding sequence or individual domains was transformed into the BL21 (DE3) bacterial cells. The single colony was inoculated in LB medium containing ampicillin and incubated at 37℃, 180rpm for overnight. Added 1% inoculum in 1L auto-induction medium. The cells were grown for 5-6 hours at 37°C, followed by 18℃ overnight at 180-200rpm. The cells were harvested by centrifuging at 6000 X g, 20 min. The cell pellet was either processed immediately or stored at -80℃. The cell pellet was resuspended in Lysis Buffer (20 mM Tris-Cl pH 8.0, 300 mM NaCl, 0.2% Triton X-100, 10% Glycerol, 1 mM Lysozyme, 1 mM PMSF, 10 mM Imidazole and 1 mM DTT). The pellet was resuspended completely and incubated on ice for 15-20 min. Sonication of the sample was done with the following conditions: 60% amplitude - 2 sec ON/6 sec OFF – 10 min. The solution was centrifuged at 12000 rpm, 4℃, 30 min. The remaining supernatant is mixed with Ni-NTA beads (Takara) and incubated on the end-to-end rotator for 3 hours, 8-10 rpm. The beads were washed 5-6 times with Wash Buffer (20 mM Tris-Cl pH 8.0, 300mM NaCl, 0.2% Triton X-100, 10% Glycerol, 20 mM Imidazole). The protein was eluted with Elution Buffer (20 mM Tris-Cl pH 8.0, 300 mM NaCl, 0.2% Triton X-100, 10% Glycerol, 100 mM Imidazole) 4 times and resolved on SDS-PAGE. The pooled fractions were dialyzed using 7-KDa cutoff dialysis bag in Dialysis Buffer I (20 mM Tris-Cl pH 8.0, 300 mM NaCl, 0.2% Triton X-100, 10% Glycerol) for 12-14 h. A buffer change to Dialysis Buffer II (20 mM Tris-Cl pH 8.0, 300 mM NaCl, 0.2% Triton X-100, 50% Glycerol) was done for 5-6 hours. The protein was collected and stored at -80℃.

2.2.5 In vitro droplet formation assay

The purified proteins were quantitated using BCA assay kit (Thermo Scientific) as per the manufacturers protocol. The droplet reactions were set up in 1.5 mL tubes as a 50 µL reaction with the following components:

1M Hepes pH 7- 22 µl

TE pH 7.5 -12 µl

BSA- 5.5 µl

60% PEG-9 µl

100 mM DTT- 1.5 µl

DNA – 200 ng

Protein: as per the concentration required

The solution upon turning turbid was visualized under the confocal microscope with channels AF 488 and DIC, and images were acquired. 1,6 hexanediol (Sigma) was added directly to the above reactions for the droplet disruption experiments and incubated for 10, 15 and 30 min before visualization.

2.2.6 Fluorescence recovery after photobleaching (FRAP) assay

The droplet formation assay described above was carried out. Upon visualization of droplets under fluorescence microscope, the points to be bleached were selected using the pointer tool in the FRAP settings of SP8 microscope (Leica). The following settings were used to image the selected points:

Imaging parameters	pre-bleach	bleach	Post-bleach PB1	Post bleach PB2
iterations	1		15-20	15-20
time/iteration	0: 20: 640 min:sec:1/1000	0:7:400	0: 20: 640 min:sec:1/1000	0: 20: 640 min:sec:1/1000

2.2.7 Databases and analyses for structure prediction

The amino acid sequences used for SATB1 and SMC1A were obtained from Uniprot database with accession numbers Q01826 and Q14683, respectively. The disorder prediction was carried out using PONDR with VL_X setting, as well as with IUPred with long-disorder option and PLAAC database for prion-like similarity.

Alphafold tool using Colab2 notebook was used to predict de novo structures of SATB1 with minimal modifications to the default settings. The predicted structure of SMC1A was also obtained from the Alphafold database. The structures predicted were visualized in the PyMOL software.

2.2.8 Western blotting

BCA method was used to quantify the proteins after the cells were lysed using RIPA lysis buffer (50 mM Tris-HCl at PH 7.4, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, 0.25% sodium deoxycholate, 1 mM PMSF, and 1X protease inhibitor cocktail (Roche). After the whole protein was separated on a 10 %SDS polyacrylamide gel, it was transferred to a Millipore PVDF membrane. Blocking of non-specific interactions was done with 5% skimmed milk. the following antibodies were used to probe the blot: anti-Actin (1:1000, Millipore), anti-SATB1 (1:1000, Santacruz Biotechnologies), and anti-Smc1a (1:1000) at 4°C overnight. Next day, multiple washes were given to the blots with 1x TBST (500mM NaCl, 20mM Tris pH 7.5, 1mM EDTA, 1% Tween 20) and anti-primary secondary antibodies tagged with HRP (anti-mouse IgG-HRP or anti-rabbit-HRP) were used to probe the blots followed by extensive washing with 1x TBST. Using the ImageQuant LAS 4000 system (GE Healthcare Life Sciences) and the ECL luminescence detection reagent (BIO-RAD), the signals were shown.

2.2.9 Site-directed mutagenesis

SATB1 GFP mutations were made using site-directed mutagenesis as described (Larson et al., 2017). Briefly, amino acid mutations that were identified targets were used for designing oligos. One or multiple bases were mutated to gain the desired amino acid change using WT SATB1-GFP in pEGFP-N1 mammalian expression vector. A two-step PCR reaction with the conditions: denaturation 98°C for 5 min, cycled for 30 cycles with 98°C for 30 sec, 68°C for 8 min, and final extension 68°C for 10 min. The PCR products were run on 1% agarose gel. The desired DNA fragment was excised and gel extracted. The DNA was treated with DpnI at 37°C for 3 hr to remove parental plasmids. The DNA was re-purified and transformed in DH5α ultra-competent bacteria. The colonies were screened single digestion with EcoRI. The positive clones were confirmed with Sanger sequencing. The oligos used for PCR are the following:

Mutation	Sequence
P181L_F	CTTGCCTCTCGAACAATGGTCGCACAC
P181L_R	TGTTCGAGAGGCAAGTCTTCTAGTTTGGGG
E530K_F	GGTTGTGCAAGCTGTTACGCTGGAAAGAAGA
E530K_R	ACAGCTTGCACAACCATCCCTGGCT
E547K_F	CCCTGTGGAAGAACCTCTCCATGATCCGAAGGT
E547K_R	GGTTCTTCCACAGGGTTCTGTTTTCTGGAGA

E407G_F	GCTTTCAGGAATCCTCCGAAAGGAAGAGGACC
E407G_R	AGGATTCCTGAAAGCAAGCCCTGAGTTCTG
Q525R_F	CAAAAGCCGGGGATGGTTGTGCGAGCTG
Q525R_R	CATCCCCGGCTTTTGGTTGCTGCAAC
N736I fs_F	AGATAAAATACTAACACCCTTTTTTCAGTGAA
N736I fs_R	GTTAGTATTTTATCTTGGACACTCTCTTCCA
K51E_F	CAGGTGCAGAAATGCAGGGAGTGCCTTTAAAAC
K51E_R	GCATTTCTGCACCTGTACTCCCAAGCC
PDZ_EF_AA_F	CCAGCACAGCTGCTGCATGCTCCTCCTTGCAAT
PDZ_EF_AA_R	GCATGCAGCAGCTGTGCTGGTGAGAAAGGATAT
KWN136-138AAA FP	CTGGAGCCGCCGCTCCAACCTGGATTAGCCCT
KWN136-138AAA RP	GTTGGAGCGGCGGCTCCAGTTCCACTGTCTTACGT
DEL575-650 FP	GCTTCCACTGAAATCACCGCGTTGCTCTCCTGTTCAT
DEL575-650 RP	GCGGTGATTTCAGTGGAAGCCTTGGGAATCCTCCA
P468E FP	CAGCACACCAGAAAGCCGTCCTCCCCAGGTGAAAACA
P468E RP	GGACGGCTTTCTGGTGTGCTGATGAGGGGGGCAGGA

2.2.10 Cell culture, transfection and Reporter assay

Growth media supplemented with 10% fetal bovine serum (Gibco/Invitrogen) was used to cultivate HEK 293T cells, which were then kept in a humidified incubator at 37°C and 5% CO2. Transfection of mutant clones was performed using Lipofectamine 3000 (Invitrogen). For the reporter assay, HEK293T cells were transfected with IL2 promoter cloned in pGL3 basic. WT SATB1 or mutants were cotransfected with Il2-pGL3b. An equal amount of Renilla expressing pTRK vector was also transfected in each well. Cells were harvested after 48 hrs and subjected to passive cell lysis. Dual luciferase assay was performed as per the manufacturer's protocol (Promega).

2.3 Results

2.3.1 Satb1 sequence harbors low-complexity or intrinsically disordered regions

Ubiquitously found biological condensates formed via LLPS have recently been implicated in diverse processes like transcription (Boija et al., 2018), chromatin organization (Ahn et al., 2021; Gibson et al., 2019), cell-cycle and splicing (Takayama and Inoue, 2022; Wagh et al., 2021). In vitro reconstituted chromatin exhibits an inherent capacity to produce a condensed yet dynamic liquid state under physiological conditions, demonstrating that LLPS may be adequate to achieve the level of compaction required for structuring the chromatin within the nucleus (Gibson et al., 2019). Furthermore, it was shown that transcriptional factor recruitment to acetylated enhancers is crucial for gene expression by forming a phase-separated structure (Cho et al., 2018); multi-bromodomain proteins such as BRD4 have the ability to produce condensates with acetylated chromatin (Sabari et al., 2018). Satb1 also harbors oligomerization property through its ULD/PDZ-like domain (Fig 2.3.1A), which has previously been shown to exist as a dimer (Galande et al., 2001) or a tetramer (Wang et al., 2012). Satb1 also displays distinct speckle-like nuclear foci in DP thymocytes (Fig 1.2.2B), a feature of condensate formation (Narlikar, 2020). We questioned whether Satb1 possesses attributes to undergo LLPS. First, in-silico analysis on Satb1 protein sequence using disorder detection tools such as IUPred and PLAAC revealed multiple regions of low-complexity or intrinsic disorder embedded in between distinct structured domains (Fig 2.3.1A and Fig 2.3.1B). This is an atypical profile as multiple other TFs known to phase transition contain a relatively contiguous stretch of intrinsically disordered region (IDR) (Fig 2.3.1C). Since the full length SATB1 structure is not yet determined, primarily due to its inability to form an organized crystal lattice (Wang et al., 2012), we confirmed the disorder in SATB1 amino acid by de novo structure prediction using Deepmind Alphafold prediction tool. We found a similar disorder pattern which displays IDRs throughout the sequence, intervened by its structured domains (Fig 2.3.1D and 2.3.1E). Furthermore, we generated de novo structure predictions for oligomeric SATB1 -dimer and trimer (Fig 2.3.1F). We observed that apart from the dimerizing ULD, SATB1 IDRs may interact with each other. Collectively, these results suggest that SATB1 harbours intrinsically disordered regions (IDRs) in its protein sequence, and could potentially undergo phase transitions via their interactions.

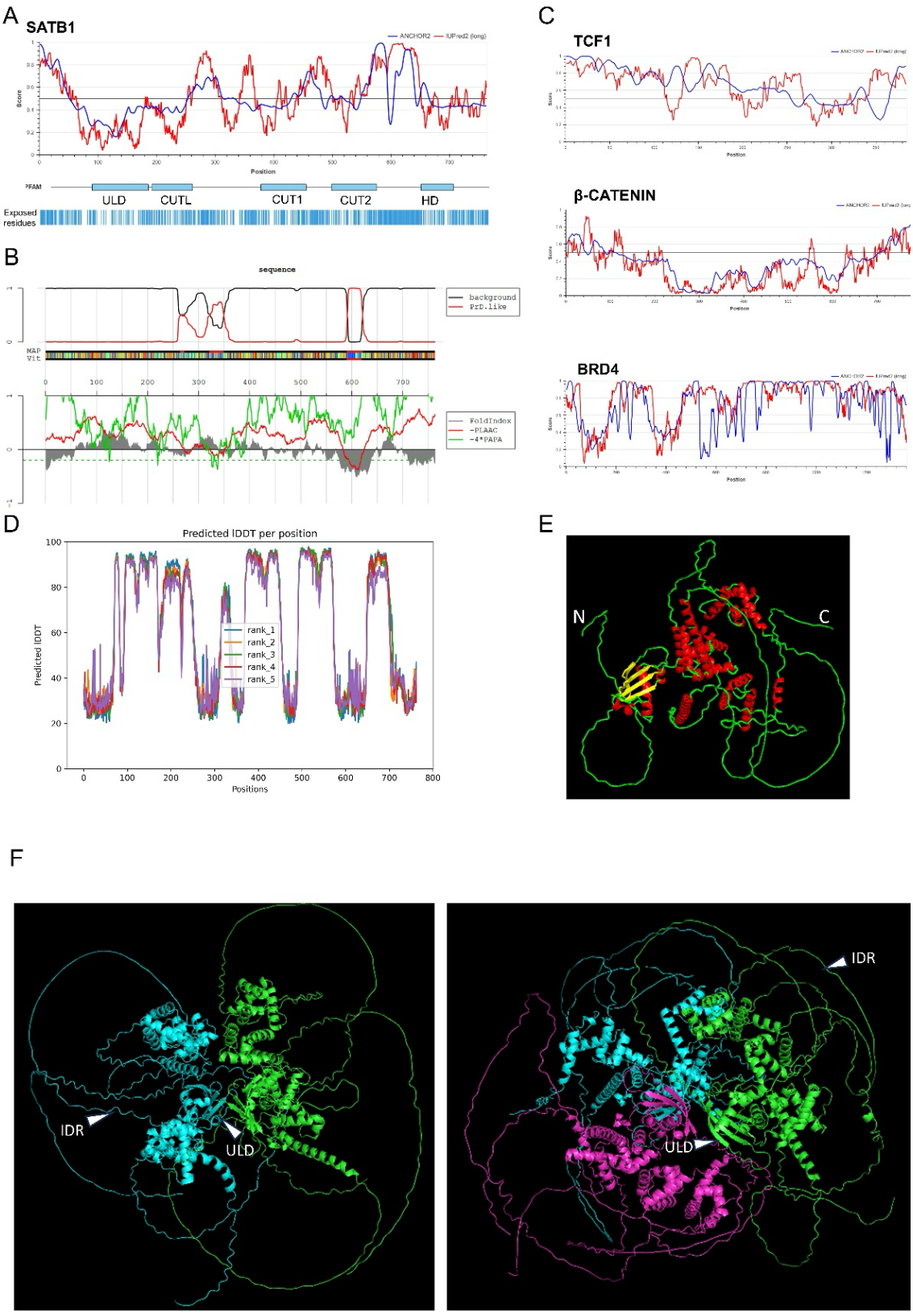


Figure 2.3.1: A In silico analysis of Satb1 sequence shows low-complexity or intrinsically disordered regions. A and **B.** Satb1 amino acid sequence was acquired from

Uniprot database (Q01826) and fed to publicly available softwares IUPred and PLAAC. Both analyses showed a high degree of disorder in Satb1 protein. Scores above the baseline indicates increased disorder. Interestingly, Satb1 disorder occurs in the linkers of all the structured functional domains. Exposed residues are also shown, which show an enrichment in the disorder regions of SATB1. **C.** Similar disorder prediction analyses on other factors known to phase separate in vitro. Comparatively, these factors, especially TCF1 and β catenin showed disorder on one end -N terminus of their respective sequences. **D.** In silico analysis of SATB1 sequence showing confidence scores for accurate structure prediction. Contrasting to IDR analysis shown in a, the higher IDDT score depicts more structured regions, and the lower scores depicts IDR-like regions in the amino acid sequence of SATB1. **E.** De novo structure prediction of monomeric SATB1 protein. It shows a high degree of disordered chains, indicated in green. **F.** A similar structure prediction analysis was done for oligomeric -dimeric and trimeric SATB1, showing that the predicted IDRs could interact with other IDRs as well as functional domains of SATB1.

2.3.2 Satb1 displays biophysical properties of liquid-liquid phase separation (LLPS) in vitro

To test the ability of Satb1 to undergo phase separation, we purified full-length Satb1 protein and performed in vitro droplet formation assay (Larson et al., 2017). Briefly, under fixed conditions such as salt concentration, pH, crowding agent and DNA concentration, we added varying amounts of protein to the mixture and observed at different time points under a fluorescence microscope. We observed that Satb1 forms liquid-like droplets in the presence of DNA (Figure 2.3.2A) and that the droplet size is proportional to the protein concentration (Fig 2.3.2B). We also find that Satb1 could also undergo phase separation in the absence of any DNA (Fig 2.3.2C), albeit with compromised droplet shape, and requires a higher critical concentration of ~0.125 μM compared to that of Satb1-DNA droplets which starts to coalesce at ~0.03 μM. Furthermore, to characterize the domains involved in condensate formation by Satb1, we purified previously described Satb1 functional domains containing using NiNTA affinity purification. We performed the droplet assay on the individual purified domains similar to that of full length Satb1. Upon visualization, we observed that both ULD/PDZ and DBD (MD+HD) were able to demix in vitro (Fig 2.3.2D and 2.3.2E). This conforms with the observation that Satb1 contains IDRs in its ULD/PDZ-like as well as DNA-binding Cut+HD regions shown in Fig 2.3.1. Although, we observed difference in droplet sizes among the two proteins assayed under identical conditions.

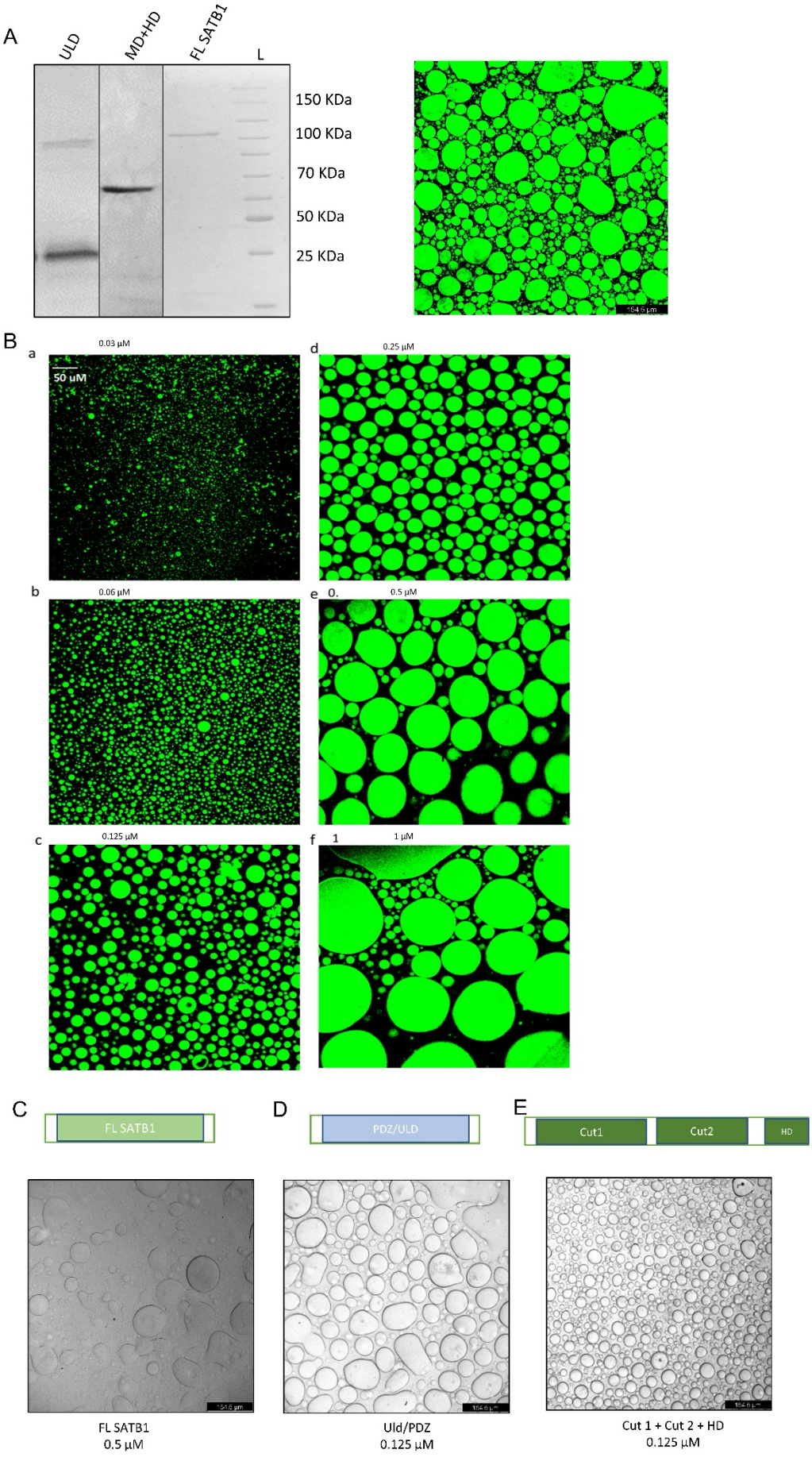

A
ULD
MD+HD
FL SATB1
L
150 KDa
100 KDa
70 KDa
50 KDa
25 KDa
B
a
0.03 µM
50 uM
d
0.25 µM
b
0.06 µM
e
0.5 µM
c
0.125 µM
f
1 µM
C
FL SATB1
D
PDZ/ULD
E
Cut1
Cut2
HD
FL SATB1
0.5 µM
Uld/PDZ
0.125 µM
Cut 1 + Cut 2 + HD
0.125 µM

Figure 2.3.2: Satb1 demonstrate the phenomenon of condensate formation in vitro. A. Full length SATB1 was purified from E.coli. In vitro droplet formation assay was performed with the purified protein. YOYO-1 (Thermo)-labelled linearized DNA (containing Satb1 target sequence) was used in the assay. The droplets containing Satb1 and bound DNA (green) were visualized. **B.** Similar to **A**, droplet formation assay was carried out but without the presence of DNA and visualized with DIC light microscope. **C.** Droplet assay with increasing concentrations of Satb1 protein was carried out, showing a directly proportional relationship between Satb1 concentration and droplet size. **D.** Protein-protein interaction domain of SATB1 -ULD/PDZ was purified and subjected to the droplet formation assay. As observed, ULD was sufficient to form the demixed condensates in vitro. **E.** Similar to **D**, MD+HD, containing the DNA binding Cut domains and sequence specifying homeodomain) also showed the property to phase separate, although with higher saturation concentration in vitro.

To assess whether Satb1 posses other characteristics of LLPS, we performed fluorescence recovery after photobleaching (FRAP) assay to ascertain molecular diffusion in the droplets (Fig 2.3.3A and Fig 2.3.3B) We observed that after FRAP (t=0), the recovery of bleached region starts after ~3 sec, and plateaus in ~25 sec, showing a recovery nearing 80% of maximum fluorescence intensity (pre-bleach). We also observed droplet fusion events during the assay, further suggesting liquid-like behaviour of Satb1 (Fig 2.3.3Aii and Fig 2.3.3Aiii). We also performed a pelletting assay with full-length Satb1, in which the protein concentration was gradually increased, keeping the DNA concentration constant (Fig 2.3.3C). We found that similar to previously characterized proteins which form condensates in vitro (Larson et al., 2017; Sabari et al., 2018), as Satb1 concentration increases, it is able to capture more DNA into its condensate in vitro. As LLPS occurs primarily via weak and reversible hydrophobic interactions (Brangwynne et al., 2009), changes in the hydrophobicity via aliphatic alcohol 1,6 hexanediol could dissolve these condensates in vitro and in vivo (Larson et al., 2017; Muzzopappa et al., 2022). We performed the droplet assay as described above. Before visualizing, we incubated the solution with 1,6 hexanediol at varying concentrations for different time points (Fig 2.3.3D). We observed that Satb1 droplets disintegrate with either increasing concentration of 1,6 hexanediol or higher incubation period. Collectively, these data indicate that Satb1 could undergo condensate formation in vitro in a concentration dependent manner. Satb1 droplets also display key liquid-like features of LLPS condensates in vitro.

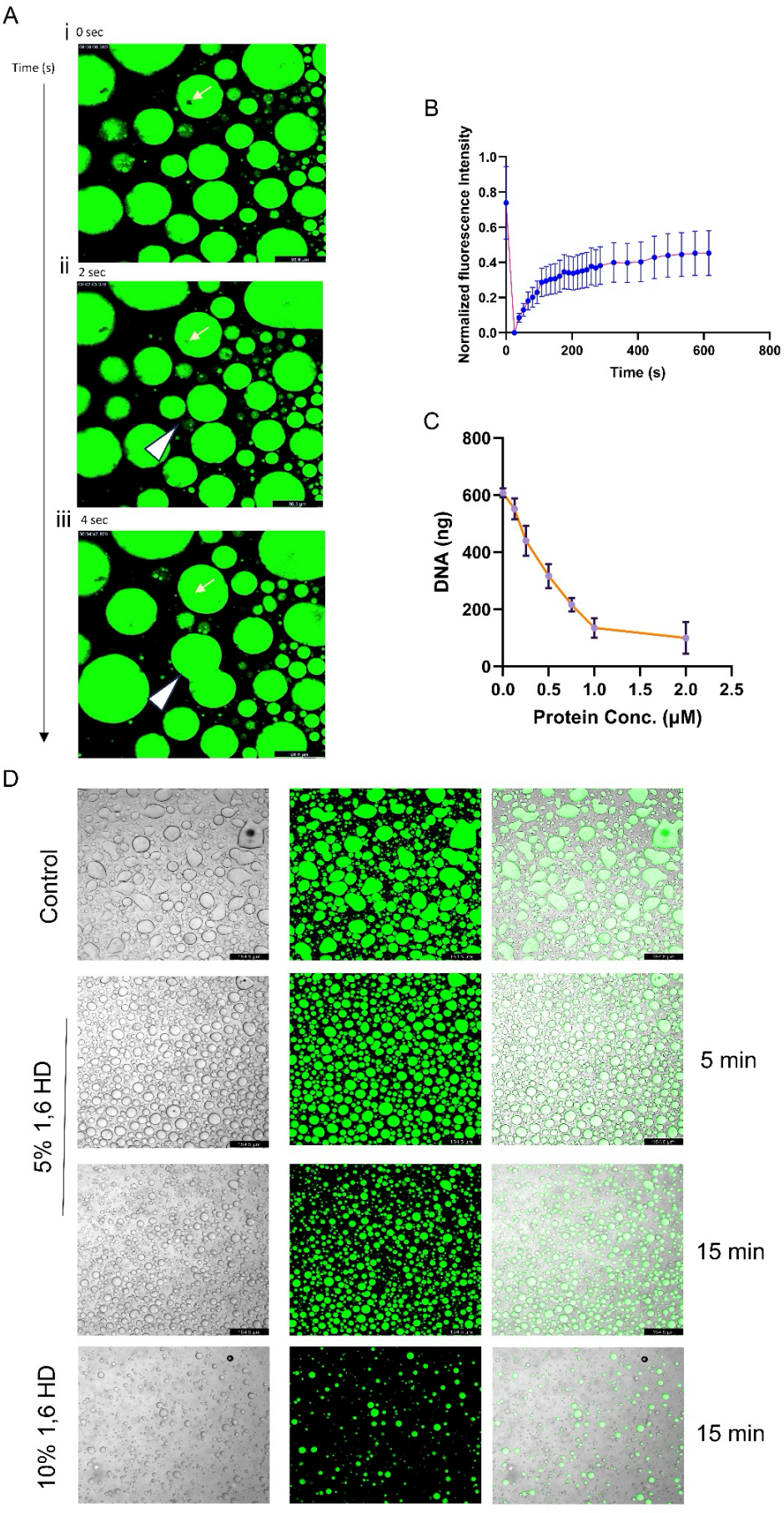

A
i 0 sec
Time (s)
ii 2 sec
iii 4 sec
B
Normalized fluorescence Intensity
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
Time (s)
C
DNA (ng)
800
600
400
200
0
0.0
0.5
1.0
1.5
2.0
2.5
Protein Conc. (μM)
D
Control
5% 1,6 HD
5 min
15 min
10% 1,6 HD
15 min

Figure 2.3.3 Satb1 displays biophysical properties of liquid-liquid phase separation (LLPS) in vitro. A. Droplet formation assay for Satb1 was performed in vitro. Fluorescence recovery after photobleaching (FRAP) was performed followed by continuous imaging upto 20 minutes. The bleach point is marked inside the droplet in (i). (ii) shows immediate recovery at the bleach site. (iii) At the bottom, marked by an arrowhead, two droplets show fushion or merging behaviour to coalesce into a bigger droplets. **B.** Quantification of the recovery in the fluorescence intensity at the fluorescence depleted point, shown in a (i), as a function of time. **C.** Droplet formation assay mixtures were centrifuged. Supernatant was collected and DNA concentration at 260 nm was measured. The plot depicts that the amount of free DNA in the 'mixed' solution is inversely proportional to the concentration of SATB1 present. **D.** Individual mixtures from the droplet formation assay were incubated with the denoted concentrations of 1,6 hexanediol (HD). The images shown here describe the effect of 1,6 HD in disrupting the weak π-π interactions in Satb1 condesates.

2.3.3 Satb1 forms dynamic nuclear foci in thymocytes

Over the past few years, several transcription factors have demosntrated the capability of undergoing liquid-liquid phase separation (LLPS) properties, contributing to their regulatory function. Previous studies have illustrated that the domain structure of Satb1, characterized by intermittent intrinsically disordered regions (IDRs), facilitates its oligomerization and formation of condensates in vitro. We then investigated whether Satb1 exhibits phase transitions in vivo. Through immunostaining and confocal imaging of sorted thymic populations, we observed distinct Satb1 foci concentrated within the nucleus (Fig 2.3.4A), with their distribution dependent on varying Satb1 concentrations across different thymic subsets. Additionally, upon staining for the Satb1 interactors Smc1a and Ctcf, we noted that Smc1a, like Satb1, forms nuclear foci, whereas Ctcf shows diffuse binding throughout the nucleus. Furthermore, Satb1 conditional knockout (cKO) DP cells exhibited a moderate yet significant reduction in Smc1a foci counts (normalized with total fluorescence intensity), as quantified (Fig 2.3.4B). To assess Smc1a's capability for condensate formation, we analyzed its amino acid sequence for disorder prediction. Our findings revealed that both coiled-coil regions of Smc1a contain low complexity sequences (Fig 2.3.4C), consistent with recent reports of IDRs occurring within coiled-coil domains (Kastano et al., 2021). 2.3.2.3.2.3. Although the predicted structure suggests minimal disorder, as the coiled-coil domains stabilize through α-helix formation (Fig 2.3.4D).

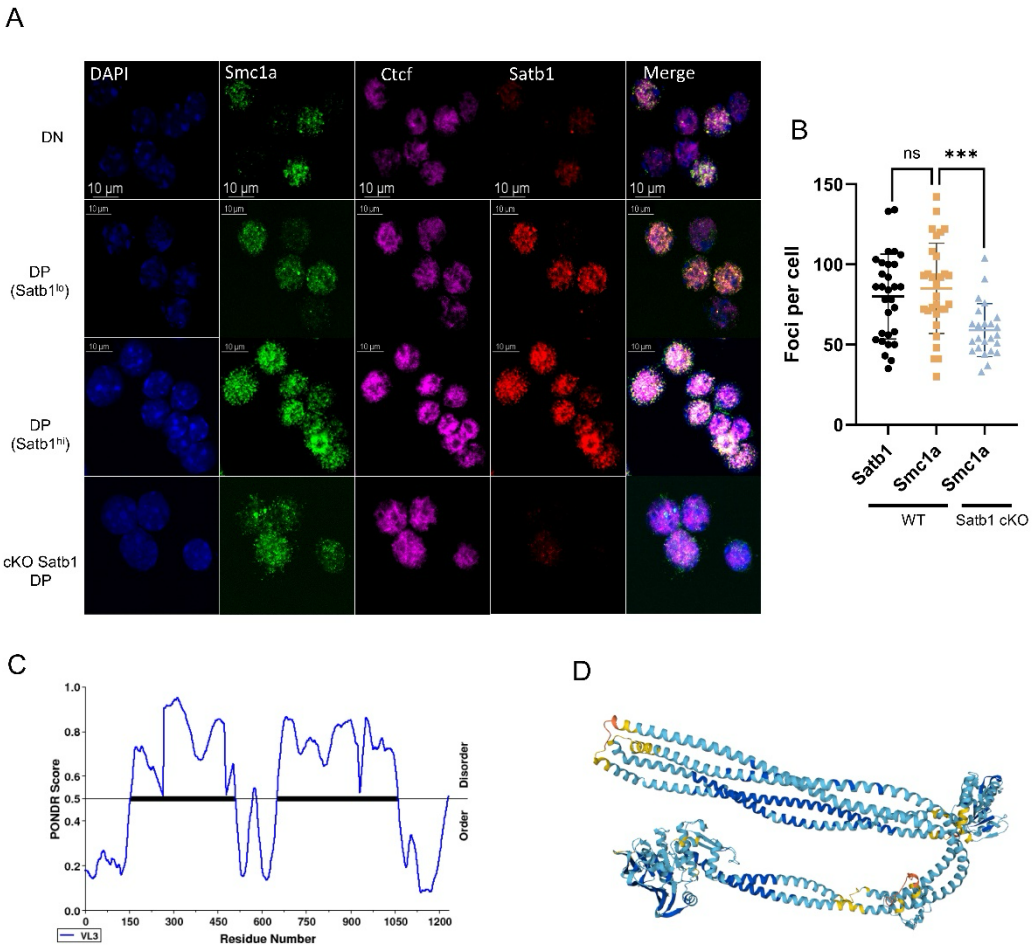


Figure 2.3.4: Satb1 forms dynamic nuclear foci in thymocytes. A. Different stages of T cell development with variable expression of Satb1 were sorted, as indicated. Immunostaning images for Satb1 (red), Smc1a (green), Ctcf (magenta) and DAPI (blue) are shown. As shown, Satb1 and Smc1a form distinct nuclear foci suggesting condensate formation, whereas Ctcf display a more diffused staining patterns. **B.** Quantitation of the number of foci formed by Smc1a in WT and Satb1 conditional knockout DP cells. The reduction in Smc1a LLPS may be indicative on the dependency on Satb1. n=3. Graphpad v10 was used for statistical analysis, two-sided ANOVA was performed *** $p<0.01$, ns non-significant. **C.** Disorder prediction of Smc1a amino acid sequence using PONDR tool, showing high degree of disorder in the coiled-coil domains. **D.** Predicted Alphafold structure for SMC1a showing minimal disordered regions.

To ascertain the LLPS nature of Satb1 condensates in developing T cells, we utilized the droplet dissolution method by 1,6 hexanediol which has been widely used to study phase separation properties of proteins *in* vitro (Larson et al., 2017; Sabari et al., 2018). We cultured sorted DP thymocytes with 5% and 10% concentrations of 1,6 hexanediol (1,6 HD) for different time periods. Upon staining the cells with Satb1, Smc1a and Ctcf we observed that treatment with 5% 1,6 HD resulted in decreased Smc1a and Satb1 foci (Fig 2.3.5A). However, treatment with 10% concentration did not reveal any significant change compared to untreated. This might be attributed to the increased chromatin condensation observed with higher concentration of 1,6 HD (Duster et al., 2021). Additionally, we performed 1,6 HD washout and cultured the DP cells before staining, in order to assess the reversal of foci disruption caused by 1,6 HD. Indeed, we observed a moderate (~20%) recovery in the foci number upon washing out 1,6 HD and quantified the same (Fig 2.3.5B). Collectively, these data demonstrate that Satb1 along with Smc1a forms nuclear condensates which colocalize, and display a liquid-like propensity.

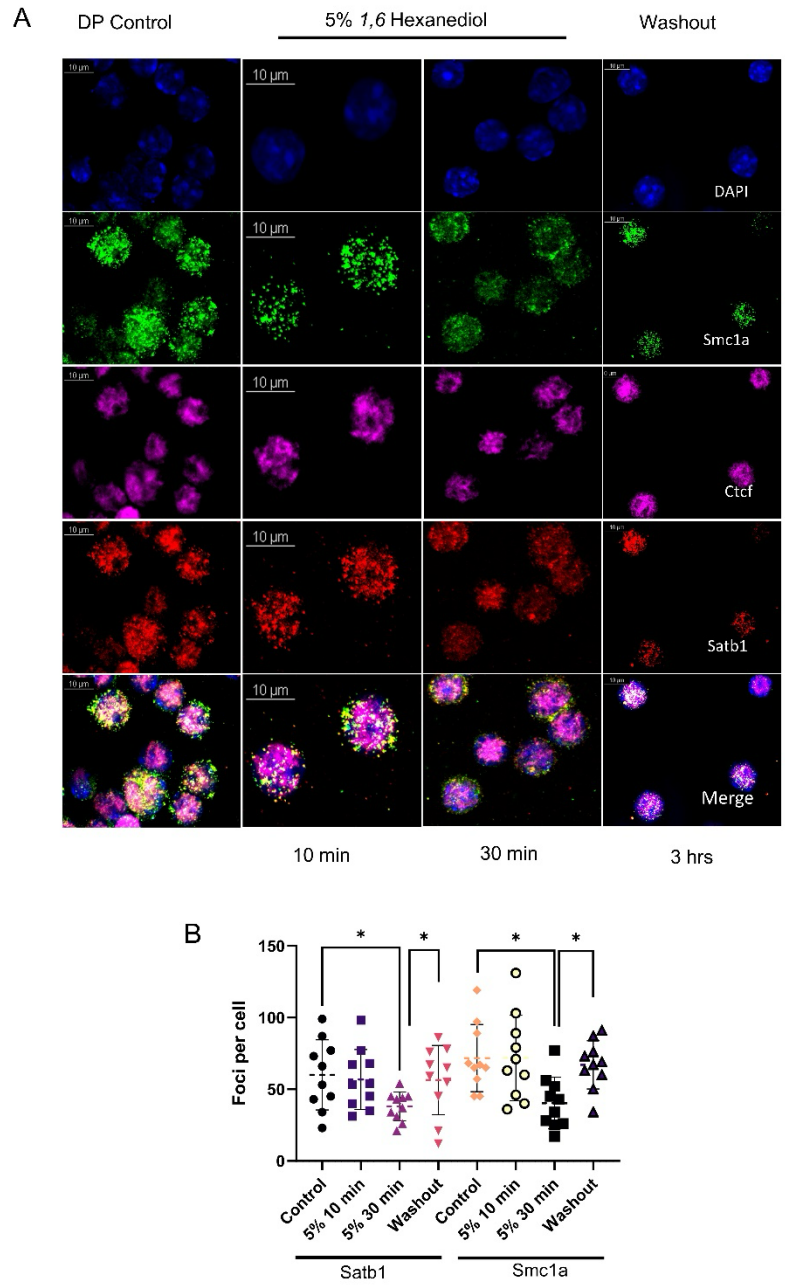


Figure 2.3.5: Satb1 condensates show liquid-like property in DP thymocytes. A. Sorted DP thymocytes were cultured in vitro with or without the presence of 1,6 hexanediol (1,6 HD). Immunostaining images shown for Satb1 (red), Smc1a (green) and Ctcf (magenta) suggest the reduction in foci formation by Satb1 and Smc1a upon 1,6 HD treatment compared to untreated control. Images after removal of 1,6 HD shows recovery in foci formation by Satb1 and Smc1a. **B.** Quantitation of the number of foci in Satb1 and Smc1a in control, 1,6 HD treated and 1,6 HD washed out samples. n=3. Graphpad v10 was used for statistical analysis, two-sided ANOVA was performed * $p<0.05$, ns non-significant.

2.3.4 Mutations in SATB1 lead to modulation in its ability to form condensates

Mutations in protein-coding sequences of thousands of genes have been identified and have been linked to disease. Since most variations occur in the intrinsically disordered regions of proteins, which show poorly defined features, their physiological impact is only recently being linked to phase separation phenomenon (Claussnitzer et al., 2020; Dyson and Wright, 2005; Tsang et al., 2020). SATB1 expression is crucial for T cell development as well as neurodevelopment (Alvarez et al., 2000). Only recently specific mutations in SATB1 coding sequence have been attributed to various neurodevelopmental disorders such as intellectual disability, dental abnormalities, epilepsy, behavioural issues and facial dysmorphisms (den Hoed et al., 2021). One study also reported a mutation in SATB1 causing its premature truncation in lung cancer (Selinger et al., 2011). Another study showed a pathogenic variant of SATB1 in a case of Trisomy X Syndrome (Necpal et al., 2021). We hypothesised whether these single nucleotide polymorphisms (SNPs) affect the condensate formation of SATB1. As SATB1 contains IDRs throughout its polypeptide sequence (Fig 2.3.6A), for assaying effect of specific mutations, we selected mutations in both structured as well as IDR regions of SATB1 as indicated (Fig 2.3.6B). We created amino acid mutations using site-specific mutagenesis using WT SATB1 cloned in a GFP containing vector. We confirmed the expression of mutant proteins via overexpression of confirmed mutant clones in HEK293T (Fig 2.3.6C). To assess the affect of the mutations on the condensate formation by SATB1, we performed live-cell imaging of HEK293T upon overexpression of these clones. We found that WT SATB1 forms distinct nuclear foci. Comparatively, K51E mutation in the IDR of the ULD domain as well as the ΔPolyQ mutant in the C terminal IDR showed stark reduction in the number of nuclear foci. Similarly, mutations in structured regions – E530K in the DNA binding CUT domain as well as mutatios EF_AA and KWN_AAA, which are disrupt dimerization of the PDZ/ULD domain – show a marked reduction in condensate formation by SATB1 (Fig 2.3.6D). While, some of the mutations assayed did not show a significant change in phase separation potential of SATB1 as shown (Fig 2.3.6D and Fig 2.3.6E). Finally, to assess the transcriptional impact of SATB1 mutations, we performed a dual luciferase reporter assay. We utilized SATB1 cognate binding site IL2 promoter (Kumar et al., 2005) activity as the readout for changes in SATB1 transactivation function. We found that the mutations which resulted in reduced condensate formation by SATB1 also resulted in reduced activity of IL2 promoter. Collectively, these results confirm that specific mutations in SATB1 can affect its phase separation potential, which in turn leads to transcriptional changes of SATB1 targets.

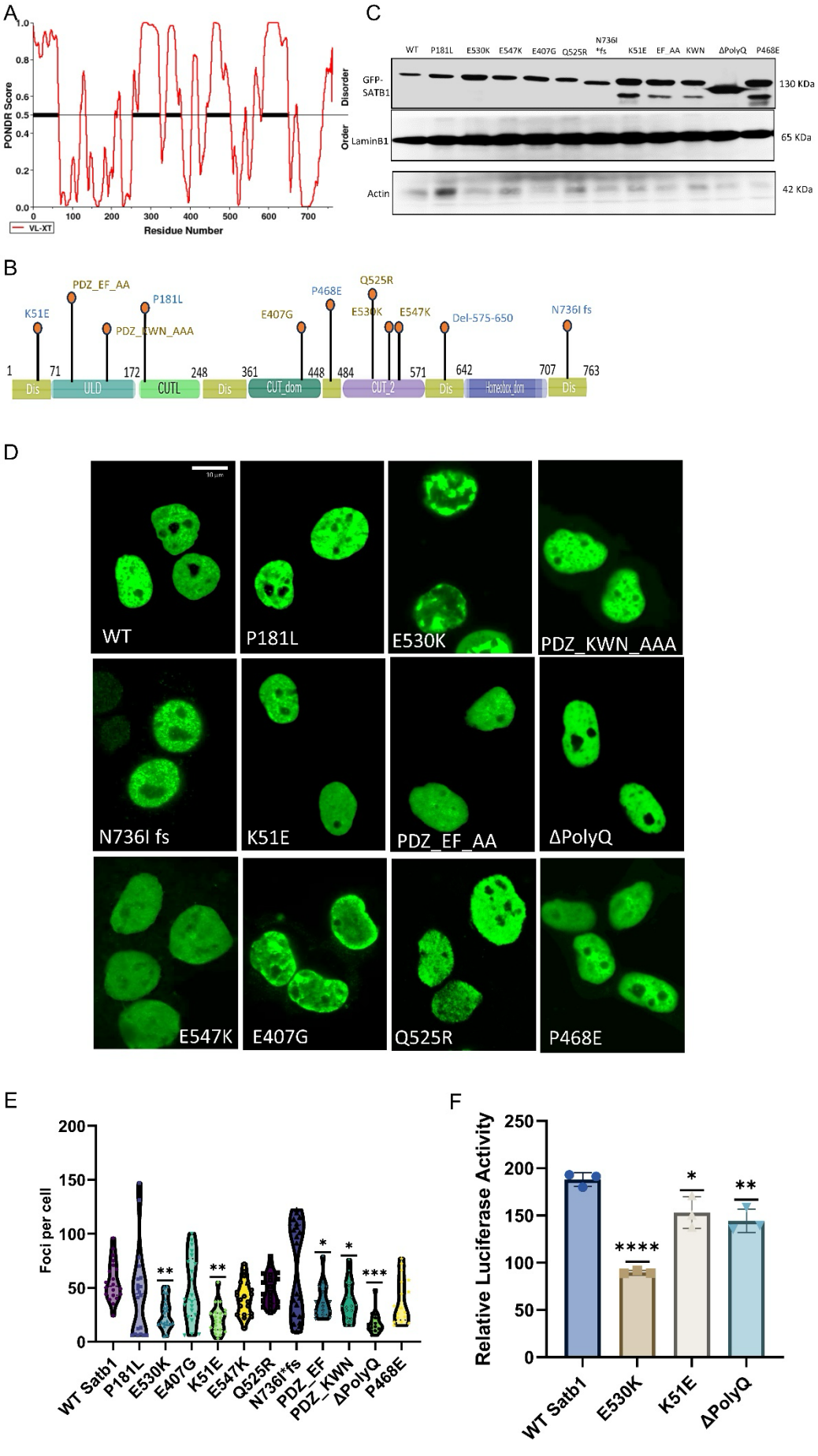

A
PONDR Score
Residue Number
VL-XT
Disorder
Order
C
WT
P181L
E530K
E547K
E407G
Q525R
N736I *fs
K51E
EF_AA
KWN
ΔPolyQ
P468E
GFP-SATB1
LaminB1
Actin
130 KDa
65 KDa
42 KDa
B
PDZ_EF_AA
K51E
P181L
PDZ_KWN_AAA
E407G
P468E
Q525R
E530K
E547K
Del-575-650
N736I fs
Dis
ULD
CUTL
CUT_dom
CUT_2
Homeobox_dom
D
WT
P181L
E530K
PDZ_KWN_AAA
N736I fs
K51E
PDZ_EF_AA
ΔPolyQ
E547K
E407G
Q525R
P468E
E
Foci per cell
WT Satb1
P181L
E530K
E407G
K51E
E547K
Q525R
N736I*fs
PDZ_EF
PDZ_KWN
ΔPolyQ
P468E
F
Relative Luciferase Activity
WT Satb1
E530K
K51E
ΔPolyQ

Figure 2.3.6: Mutations in SATB1 lead to modulation in its ability to form condensates. **A.** Disorder in SATB1 was calculated using PONDR VL-XT scoring method, and the disorder versus order is shown. **B.** Schematic of SATB1 domain map with the known domains. The disordered regions identified are labelled 'Dis' which are both at N -and C-termini as well as in all the linker chains. Specific SATB1 mutations causing neuropathic disorders were identified in both structured and unstructured regions as shown. **C.** Immunoblotting shows SATB1 protein expression for all the mutant clones generated via site-directed mutagenesis, confirming that none of the mutations or clones selected lead to degradation of SATB1 upon overexpression. **D.** Live-cell imaging of SATB1-GFP WT and mutants in HEK293T. The imaging shows that all in all the mutations SATB1 remained predominantly nuclear. Nuclear condensates formed by SATB1 are shown in each mutation by representative images. **E.** Quantitation of the number of foci formed by SATB1 mutants is shown. n=3, N=30. **F.** Dual luciferase assay was performed upon overexpression of SATB1 WT and mutants along with the reporter plasmid containing Il2 promoter. Luciferase activity was assayed 48 hrs post-transfection and normalized to Renilla activity (pTRK), n=4. Graphpad v10 was used for statistical analysis, two-sided ANOVA was performed * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

2.4 Discussion

Considerable research has been done on cellular organelles that are non-membrane-bound, such P granules and nucleoli (Banani et al., 2017). Recent evidences show a role of condensate formation in genome organization (Gibson et al., 2019) as well as transcriptional activation (Boija et al., 2018). In this chapter we describe the biophysical phenomenon of condensate formation by Satb1. By analyzing the amino acid sequence of Satb1 we found that SATB1 harbours intrinsically disordered regions (IDRs), which are low-complexity sequence rich; implying that these segments of SATB1 protein are unable to form a structured domain whose functions could be attributed to previously characterized protein folds. We further performed de novo structure prediction of monomeric and oligomeric forms of SATB1 showing that a major proportion of SATB1 amino acid composition is intrisically disordered. The intervening IDRs may provide flexibility to SATB1 to go in -and out of condensates with different constituents. Furthermore, SATB1 dimeric and trimeric structures show that apart from ULD, the disordered arms dangling from each monomer could interact with IDRs or structured domains of the next monomer, bolstering our hypothesis of oligomerization of SATB1 via IDRs. To validate our in silico analyses that SATB1 might be able to form demixed condensates, we used purified full-length SATB1 protein to perform droplet formation assay in vitro. Our analyses show that SATB1 forms distinct droplets which show biophysical characteristics pertaining to a liquid-like state. Nucleic acids such as double stranded (ds) or single stranded (ss)-DNA, RNA or DNA:RNA hybrid could promote phase separation by various transcription factors in vitro and in vivo (Bracha et al., 2018; Narlikar, 2020).

Our data suggests that condensate formation by SATB1 is also promoted by addition of DNA to the solution. Without the presence of DNA, SATB1 forms condensate at a relatively higher concentration and the droplets attain a ‘non-round’ or staggered morphology immediately after the assay is set up. Thus, DNA binding although dispensable, but may be important for appropriate functioning of SATB1 condensates. It was also shown that DNA incorporation into the droplets formed by SATB1 is a function of protein concentration in vitro. Upon further probing into biophysical attributes of SATB1 condensates, we found that SATB1 droplets show molecular diffusion as well as droplet fusion properties. Alcohol 1,6-hexanediol is a popular controller for dissolving LLPS condensates in phase separation investigations used in a range studies both in vivo and in vitro (Duster et al., 2021). We found that 1,6 HD could disperse SATB1 droplets in vitro in a concentration or time dependent manner. Together, these properties indicate toward a liquid-like nature of SATB1 condensates in vitro.

The complex of cohesin and CTCF proteins regulates chromatin interconnections in the genome (Sanborn et al., 2015). To ensure to preservation the higher-order chromatin architecture, these two proteins regulate the chromatin contacts in spatially segregated areas known as topologically associated domains (TADs). Throughout development, TAD formation are both crucial and consistent in eukaryotes throughout all cell types. Recently, condensate formation is gaining more attention in the field of genome organization as the missing link between form and function.

SATB1 functions as a docking site for many chromatin modifiers and nucleosome remodelers, and controls the expression of its target genes (Yasui et al., 2002). By interacting with and recruiting the SWI/SNF complexes on targeted genomic location, SATB1 controls the expression of a large number of genes (Li et al., 2018). SATB1 binds the MARs to the nuclear matrix, which coordinates the looping of Th2 cytokine genes (Cai et al., 2006) and the MHC-I locus (Kumar et al., 2007).

SATB1 appears to be involved in the dynamic structuring of open chromatin in the thymocyte nucleus (Galande et al., 2007). SATB1's N-terminal PDZ-like domain regulates it's interactions with other factors and is essential for homodimerization (Galande et al., 2001; Galande et al., 2007; Wang et al., 2012). We tested whether SATB1 could form biological condensates in vivo. Our immunostaining data suggest that SATB1 forms distinct nuclear foci. Simultaneous staining with SMC1A and CTCF showed that SMC1a/Cohesin but not CTCF form phase separated nuclear foci which co-exist with those of SATB1. Further assessment of these nuclear foci revealed that both SATB1 and SMC1a condensates are liquid like during T cell development. Upon SATB1 depletion from DP thymocytes, we observe a reduced foci formation by SMC1A, suggesting that the heterotypic interaction of SATB1 and SMC1A might promote condensate formation at the specific genomic loci which are shared by the two proteins.

Liquid-liquid phase separation (LLPS) has been linked to various diseases. Numerous studies have shown the importance of specific amino acid residues in driving phase separation mediated homeostasis, and its imbalance upon mutation, underpinning the disease state (Ahn et al., 2021; Tsang et al., 2020). For example, most of the amyotrophic lateral sclerosis (ALS)-related mutations, primarily in its helix, decrease LLPS property, and cause neurodegeneration (Conicella et al., 2016). Also, LLPS is increased by Parkinson's Disease-associated mutations (E46K, A53T) as well as by phosphorylation at amino acid S129 (Ray et al., 2020). It is yet unclear how certain mutations affect the way proteins behave during phase separation.Mutations leading to lowering of saturation concentration, on one hand, may confine the protein in a condensed system and promote the gradual shift from dynamic liquid

state to a permanent aggregation state, and vice versa. Thus, we asked if known disease mutations in SATB1 could modulate its phase separation property, perhaps by changing its saturation concentration. We used literature search as well as publicly available databases such as Cosmic Cancer database to identify specific residues of SATB1 involved in the disease. Our mutagenesis analysis revealed that mutations in IDRs as well as DNA and protein binding structured domains could reduce LLPS of SATB1. This indicated to us the importance of homotypic interaction by ULD of SATB1, as well as the weaker π-π interactions by the IDRs could interplay in forming SATB1 condensates. It was also found that reduction in SATB1 LLPS directly affected its target promoter activation in vitro, suggesting a role of SATB1 phase separation either in its localization or its recruitment of co-activators to the regulatory regions.

In conclusion, we found that Satb1 posseses the ability to form biological condensates. Furthermore, the condensates formed by Satb1 exhibit characteristic properties of liquid-like (LLPS) such as concentration dependence, molecular diffusion, droplet fusion and 1,6 hexandiol sensitivity. We found that Satb1 forms nuclear foci in thymocytes which demonstrates its condensate formation property in vivo. Structural protein Smc1a also showed foci formation in thymic nuclei, which coincided with Satb1 foci, potentially as a consequence of their heterotypic interaction. The foci formation of Smc1a is partially dependent on Satb1. We lastly found that specific amino acid residues both in Satb1 IDRs and specific domains are important in maintaining the LLPS propensity of SATB1 which may have a direct ramification on the transcription of SATB1 target genes.

2.5 References

Ahn, J.H., Davis, E.S., Daugird, T.A., Zhao, S., Quiroga, I.Y., Uryu, H., Li, J., Storey, A.J., Tsai, Y.H., Keeley, D.P*., et al.* (2021). Phase separation drives aberrant chromatin looping and cancer development. Nature *595*, 591-595.

Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev *14*, 521-535.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol *18*, 285-298.

Berry, J., Brangwynne, C.P., and Haataja, M. (2018). Physical principles of intracellular organization via active and passive phase transitions. Rep Prog Phys *81*, 046601.

Black, J.C., Choi, J.E., Lombardo, S.R., and Carey, M. (2006). A mechanism for coordinating chromatin modification and preinitiation complex assembly. Mol Cell *23*, 809-818.

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P*., et al.* (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat Struct Mol Biol *25*, 833-840.

Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M*., et al.* (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842-1855 e1816.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat Rev Genet *17*, 772.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A*., et al.* (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572 e524.

Bracha, D., Walls, M.T., Wei, M.T., Zhu, L., Kurian, M., Avalos, J.L., Toettcher, J.E., and Brangwynne, C.P. (2018). Mapping Local and Global Liquid Phase Behavior in Living Cells Using Photo-Oligomerizable Seeds. Cell *175*, 1467-1480 e1413.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

Buratowski, S., Hahn, S., Sharp, P.A., and Guarente, L. (1988). Function of a yeast TATA element-binding protein in a mammalian transcription system. Nature *334*, 37-42.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.

Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, II (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412-415.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X.*, et al.* (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*.

Claussnitzer, M., Cho, J.H., Collins, R., Cox, N.J., Dermitzakis, E.T., Hurles, M.E., Kathiresan, S., Kenny, E.E., Lindgren, C.M., MacArthur, D.G.*, et al.* (2020). A brief history of human disease genetics. Nature *577*, 179-189.

Conicella, A.E., Zerze, G.H., Mittal, J., and Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure *24*, 1537-1549.

Cremer, M., Grasser, F., Lanctot, C., Muller, S., Neusser, M., Zinner, R., Solovei, I., and Cremer, T. (2008). Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. Methods Mol Biol *463*, 205-239.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M.*, et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.

den Hoed, J., de Boer, E., Voisin, N., Dingemans, A.J.M., Guex, N., Wiel, L., Nellaker, C., Amudhavalli, S.M., Banka, S., Bena, F.S.*, et al.* (2021). Mutation-specific pathophysiological mechanisms define different neurodevelopmental disorders associated with SATB1 dysfunction. Am J Hum Genet *108*, 346-356.

Di Pierro, M., Potoyan, D.A., Wolynes, P.G., and Onuchic, J.N. (2018). Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. Proc Natl Acad Sci U S A *115*, 7753-7758.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schujiers, J., Lee, T.I., Zhao, K.*, et al.* (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374-387.

Dumetz, A.C., Chockla, A.M., Kaler, E.W., and Lenhoff, A.M. (2008). Protein phase behavior in aqueous solutions: crystallization, liquid-liquid phase separation, gels, and aggregates. Biophys J *94*, 570-583.

Duster, R., Kaltheuner, I.H., Schmitz, M., and Geyer, M. (2021). 1,6-Hexanediol, commonly used to dissolve liquid-liquid phase separated condensates, directly impairs kinase and phosphatase activities. J Biol Chem *296*, 100260.

Dyson, H.J., and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol *6*, 197-208.

Feric, M., and Misteli, T. (2021). Phase separation in genome organization across evolution. Trends Cell Biol *31*, 671-685.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686-1697.

Finn, E.H., Pegoraro, G., Brandao, H.B., Valton, A.L., Oomen, M.E., Dekker, J., Mirny, L., and Misteli, T. (2019). Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. Cell *176*, 1502-1515 e1510.

Fraser, J., Williamson, I., Bickmore, W.A., and Dostie, J. (2015). An Overview of Genome Organization and How We Got There: from FISH to Hi-C. Microbiol Mol Biol Rev *79*, 347-372.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.

Galande, S., Dickinson, L.A., Mian, I.S., Sikorska, M., and Kohwi-Shigematsu, T. (2001). SATB1 cleavage by caspase 6 disrupts PDZ domain-mediated dimerization, causing detachment from chromatin early in T-cell apoptosis. Mol Cell Biol *21*, 5591-5604.

Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev *17*, 408-414.

Gao, Z., Zhang, W., Chang, R., Zhang, S., Yang, G., and Zhao, G. (2021). Liquid-Liquid Phase Separation: Unraveling the Enigma of Biomolecular Condensates in Microbial Cells. Front Microbiol *12*, 751880.

Garcia, D.A., Johnson, T.A., Presman, D.M., Fettweis, G., Wagh, K., Rinaldi, L., Stavreva, D.A., Paakinaho, V., Jensen, R.A.M., Mandrup, S.*, et al.* (2021). An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors. Mol Cell *81*, 1484-1498 e1486.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470-484 e421.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B.*, et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614.

Handwerger, K.E., Cordero, J.A., and Gall, J.G. (2005). Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. Mol Biol Cell *16*, 202-211.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13-23.

Hu, G., Cui, K., Fang, D., Hirose, S., Wang, X., Wangsa, D., Jin, W., Ried, T., Liu, P., Zhu, J.*, et al.* (2018). Transformation of Accessible Chromatin and 3D Nucleome Underlies Lineage Commitment of Early T Cells. Immunity *48*, 227-242 e228.

Hyman, A.A., Weber, C.A., and Julicher, F. (2014). Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol *30*, 39-58.

Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Varnai, C., Thiecke, M.J.*, et al.* (2016). Lineage-Specific Genome Architecture

Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell *167*, 1369-1384 e1319.

Johanson, T.M., Lun, A.T.L., Coughlan, H.D., Tan, T., Smyth, G.K., Nutt, S.L., and Allan, R.S. (2018). Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. Nat Immunol *19*, 1257-1264.

Kao, C.C., Lieberman, P.M., Schmidt, M.C., Zhou, Q., Pei, R., and Berk, A.J. (1990). Cloning of a transcriptionally active human TATA binding factor. Science *248*, 1646-1650.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J.*, et al.* (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753-767.

Kieffer-Kwon, K.R., Tang, Z., Mathe, E., Qian, J., Sung, M.H., Li, G., Resch, W., Baek, S., Pruett, N., Grontved, L.*, et al.* (2013). Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. Cell *155*, 1507-1520.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galande, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol *9*, 45-56.

Kumar, P.P., Purbey, P.K., Ravi, D.S., Mitra, D., and Galande, S. (2005). Displacement of SATB1-bound histone deacetylase 1 corepressor by the human immunodeficiency virus type 1 transactivator induces expression of interleukin-2 and its receptor in T cells. Mol Cell Biol *25*, 1620-1633.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature *547*, 236-240.

Li, Y., Wang, J., Yu, M., Wang, Y., Zhang, H., Yin, J., Li, Z., Li, T., Yan, H., Li, F.*, et al.* (2018). SNF5 deficiency induces apoptosis resistance by repressing SATB1 expression in Sezary syndrome. Leuk Lymphoma *59*, 2405-2413.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Maeshima, K., Ide, S., Hibino, K., and Sasai, M. (2016). Liquid-like behavior of chromatin. Curr Opin Genet Dev *37*, 36-45.

Malik, S., and Roeder, R.G. (2023). Regulation of the RNA polymerase II pre-initiation complex by its associated coactivators. Nat Rev Genet *24*, 767-782.

Mitrea, D.M., and Kriwacki, R.W. (2016). Phase separation in biology; functional organization of a higher order. Cell Commun Signal *14*, 1.

Monneron, A., and Bernhard, W. (1969). Fine structural organization of the interphase nucleus in some mammalian cells. J Ultrastruct Res *27*, 266-288.

Muzzopappa, F., Hummert, J., Anfossi, M., Tashev, S.A., Herten, D.P., and Erdel, F. (2022). Detecting and quantifying liquid-liquid phase separation in living cells by model-free calibrated half-bleaching. Nat Commun *13*, 7787.

Nair, S.J., Yang, L., Meluzzi, D., Oh, S., Yang, F., Friedman, M.J., Wang, S., Suter, T., Alshareedah, I., Gamliel, A.*, et al.* (2019). Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. Nat Struct Mol Biol *26*, 193-203.

Narlikar, G.J. (2020). Phase-separation in chromatin organization. J Biosci *45*.

Necpal, J., Zech, M., Winkelmann, J., and Jech, R. (2021). Trisomy X syndrome with dystonia and a pathogenic SATB1 variant. Neurol Sci *42*, 3883-3884.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N., and Mirny, L.A. (2018). Chromatin organization by an interplay of loop extrusion and compartmental segregation. Proc Natl Acad Sci U S A *115*, E6697-E6706.

Pekowska, A., Klaus, B., Xiang, W., Severino, J., Daigle, N., Klein, F.A., Oles, M., Casellas, R., Ellenberg, J., Steinmetz, L.M.*, et al.* (2018). Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. Cell Syst *7*, 482-495 e410.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Ray, S., Singh, N., Kumar, R., Patel, K., Pandey, S., Datta, D., Mahato, J., Panigrahi, R., Navalkar, A., Mehra, S.*, et al.* (2020). alpha-Synuclein aggregation nucleates through liquid-liquid phase separation. Nat Chem *12*, 705-716.

Richter, W.F., Nayak, S., Iwasa, J., and Taatjes, D.J. (2022). The Mediator complex as a master regulator of transcription by RNA polymerase II. Nat Rev Mol Cell Biol *23*, 732-749.

Roeder, R.G. (1998). Role of general and gene-specific cofactors in the regulation of eukaryotic transcription. Cold Spring Harb Symp Quant Biol *63*, 201-218.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C.*, et al.* (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*.

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J.*, et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A *112*, E6456-6465.

Schier, A.C., and Taatjes, D.J. (2020). Structure and mechanism of the RNA polymerase II transcription machinery. Genes Dev *34*, 465-488.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Selinger, C.I., Cooper, W.A., Al-Sohaily, S., Mladenova, D.N., Pangon, L., Kennedy, C.W., McCaughan, B.C., Stirzaker, C., and Kohonen-Corish, M.R. (2011). Loss of special AT-rich binding protein 1 expression is a marker of poor survival in lung cancer. J Thorac Oncol *6*, 1179-1189.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shaban, H.A., Barth, R., and Bystricky, K. (2018). Formation of correlated chromatin domains at nanoscale dynamic resolution during transcription. Nucleic Acids Res *46*, e77.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V.*, et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Shin, Y., Chang, Y.C., Lee, D.S.W., Berry, J., Sanders, D.W., Ronceray, P., Wingreen, N.S., Haataja, M., and Brangwynne, C.P. (2018). Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. Cell *175*, 1481-1491 e1413.

Stadhouders, R., Filion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. Nature *569*, 345-354.

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y*., et al.* (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat Genet *50*, 238-249.

Sun, F., Chronis, C., Kronenberg, M., Chen, X.F., Su, T., Lay, F.D., Plath, K., Kurdistani, S.K., and Carey, M.F. (2019). Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. Mol Cell *73*, 250-263 e255.

Takayama, K.I., and Inoue, S. (2022). Targeting phase separation on enhancers induced by transcription factor complex formations as a new strategy for treating drug-resistant cancers. Front Oncol *12*, 1024600.

Thomas, M.C., and Chiang, C.M. (2006). The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol *41*, 105-178.

Tsang, B., Pritisanac, I., Scherer, S.W., Moses, A.M., and Forman-Kay, J.D. (2020). Phase Separation as a Missing Mechanism for Interpretation of Disease Mutations. Cell *183*, 1742-1756.

Wagh, K., Garcia, D.A., and Upadhyaya, A. (2021). Phase separation in transcription factor dynamics and chromatin organization. Curr Opin Struct Biol *71*, 148-155.

Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D*., et al.* (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell *174*, 688-699 e616.

Wang, S., Su, J.H., Beliveau, B.J., Bintu, B., Moffitt, J.R., Wu, C.T., and Zhuang, X. (2016). Spatial organization of chromatin domains and compartments in single chromosomes. Science *353*, 598-602.

Wang, W., Chandra, A., Goldman, N., Yoon, S., Ferrari, E.K., Nguyen, S.C., Joyce, E.F., and Vahedi, G. (2022). TCF-1 promotes chromatin interactions across topologically associating domains in T cell progenitors. Nat Immunol *23*, 1052-1062.

Wang, Z., Yang, X., Chu, X., Zhang, J., Zhou, H., Shen, Y., and Long, J. (2012). The structural basis for the oligomerization of the N-terminal domain of SATB1. Nucleic Acids Res *40*, 4193-4202.

Wei, M.T., Chang, Y.C., Shimobayashi, S.F., Shin, Y., Strom, A.R., and Brangwynne, C.P. (2020). Nucleated transcriptional condensates amplify gene expression. Nat Cell Biol *22*, 1187-1196.

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 e1528.

Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. Cold Spring Harb Perspect Biol *2*, a000596.

Zhang, Y., McCord, R.P., Ho, Y.J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell *148*, 908-921.

Zhang, Y., Wong, C.H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E.*, et al.* (2013). Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature *504*, 306-310.

Zheng, H., and Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol *20*, 535-550.

Chapter 3: Role of Satb1 in regulating TCR-dependent genes in conjunction with NFAT family proteins during T cell development

3.1 Introduction

T cell development initiates when the common lymphoid progenitors (CLPs) migrate from the bone marrow (BM) to thymus. Thymus is a dedicated organ for the development and education of T cells. The CLPs which transit to the thymus, being at initial ETP or double negative (DN1) stage not only have T-lymphocyte but B cell and myeloid potentials as well (Bell and Bhandoola, 2008; Wada et al., 2008). In the thymus, the progenitors are exposed to Notch signaling and gets committed to the T-cell fate (Rothenberg, 2014), shutting off their potential to differentiate into other lineages. Specifically, this commitment happens at the DN2a to DN2b stages (Masuda et al., 2007; Yui et al., 2010). Further, rearrangement of the T cell receptor (TCR) β locus occurs which upon success passes the β selection checkpoint–restricting cells with unproductive TCR β chain and restricts their entry into the double positive (DP) stage by inducing apoptosis (Koltsova et al., 2007; Krangel, 2009). As the thymocytes with productive TCRβ mature into DP cells, rearrangement of the TCRα chain begins. After the formation of completely rearranged TCRαβ dimer, two steps of selection occur, namely positive and negative selection (Carpenter and Bosselut, 2010; Germain, 2002; Starr et al., 2003). DP thymocytes with successful TCRαβ rearrangement undergo a series of interactions with self-peptide:MHC complexes. DP cells with moderate affinity are positively selected. Positive selection provides a survival signal to the DP cells which can then undergo CD4/CD8 fate choice (Carpenter and Bosselut, 2010). DPs with high affinity to self p:MHC are marked auto-reactive and undergo apoptosis (Murphy et al., 1990; Swat et al., 1991; Vasquez et al., 1992). This selection process is termed negative selection and is important for evading auto-immunity (Kappler et al., 1987; Palmer, 2003; Swat et al., 1991). Later development of DP cells to CD4+ or CD8+ SP cells warrants turning off one of these co-receptors, which is deemed to be one of the function of TCR signal duration and strength, at least in part (Singer et al., 2008). DP thymocytes are very sensitive to the MHC:peptide complex, much more than peripheral effector cells (Carpenter and Bosselut, 2010; Singer et al., 2008), however, the signaling cascades are much of the same (Janeway, 1992), acting in feedback and feedforward loops (Carpenter and Bosselut, 2010). Early seminal studies on TCR:pMHC interactions asked; how the positively selected thymocytes are able to avoid negative selection despite continuous exposure to p:MHC. It was found that it is the expression of different lineage specific

transcription factors and ‘fine-tune’ regulators that mediate this distinction of response to TCR signaling (Fu et al., 2014; Gascoigne et al., 2016) which leads these cells to maturation.

3.1.1 Classical and current models for TCR signal mediated fate choice

As described above, upon expression of a functional TCRαβ surface expression, the DP cell interact with the self-pMHC, and undergo positive selection. Almost at the same time most cells which are positively selected undergo a lineage fate choice to either become $CD8^+$ or $CD4^+$ SP cells. The mechanisms for fate choice have been widely debated (Singer et al., 2008).

3.1.1.1 Classical models for CD4/CD8 choice

CD4/CD8 lineage decision was originally thought to depend on the transcriptional end of either of the co-receptor genes; CD4 and CD8, as a result of a single TCR signaling cascade that enables positive selection. This was because DP thymocytes have bipotency to display both the co-receptors (Leung et al., 2001; Robey et al., 1994). According to the CD4/CD8 fate determination by the stochastic model, the gene expression of either co-receptor terminates in a 'random' manner following positive selection of DP thymocytes (Itano et al., 1994) (Chan et al., 1993; Davis et al., 1993). Following the positive selection, there is a subsequent TCR-dependent "salvation" process that ensures only SP thymocytes with compatible TCRs and co-receptors survived (Chan et al., 1994) and develop into mature T cells (Figure 3.1.1). Experimental evidences show contradictory outcomes to the stochastic model; the efficiency of lineage determined cells should be around 50% since it postulates the choice random event. But the process was shown to be highly efficient approaching up to 90% (Itano and Robey, 2000).

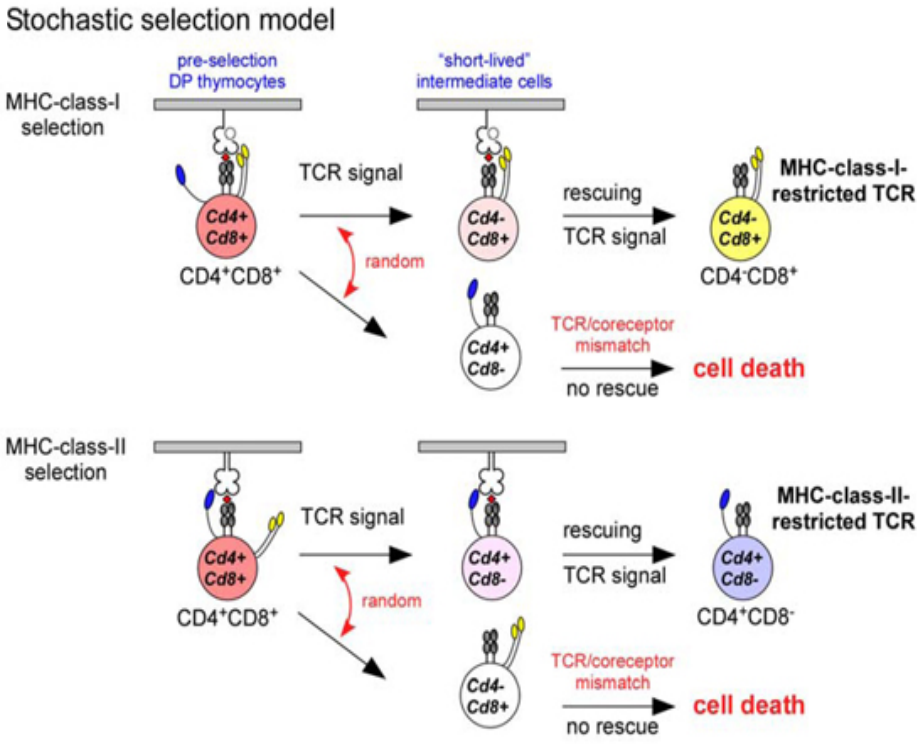


Figure 3.1.1: Stochastic model of differentiation. According to the stochastic model of lineage choice, the TCR signals arbitrarily stop the expression of either CD4 or CD8 co-receptor molecules triggering the production of 'transient' intermediate cells, which necessarily requires 'saving' by a second TCR signal, otherwise they undergo apoptosis. Due to its stochastic nature, roughly 50% of the cells should get positively selected. Reproduced from Singer et al., 2008.

According to the strength of signal model, the variations in the strength of TCR signal transduction and differences in co-receptor interactions during positive selection guides CD4/CD8 lineage choice of DP thymocytes (Itano et al., 1996). Based on the intensity or strength of the signal generated by the TCR and either CD4 or CD8 co-receptor simultaneous engagement, thymocytes are stimulated to precisely stop either Cd8 or Cd4 gene production (Itano et al., 1996). TCR and CD4 co-engagement produces a stronger signal because the cytoplasmic tail of CD4 co-receptor interacts strongly with tyrosine kinase Lck (which initiates the downstream TCR signaling events) compared to the CD8 chain (Shaw et al., 1989; Wiest et al., 1993). In contrast, TCR and CD8 co-engagement produces weaker signals (Figure 3.1.2). Later studies showed contradictory experimental evidences to this model, suggesting alternative or additional mechanisms involved in the lineage fate choice (Bosselut et al., 2001;

Erman et al., 2006; Holst et al., 2008; Love et al., 2000). Although, it explains why there is a higher number of CD4+ cells selected compared to CD8+ SP thymocytes.

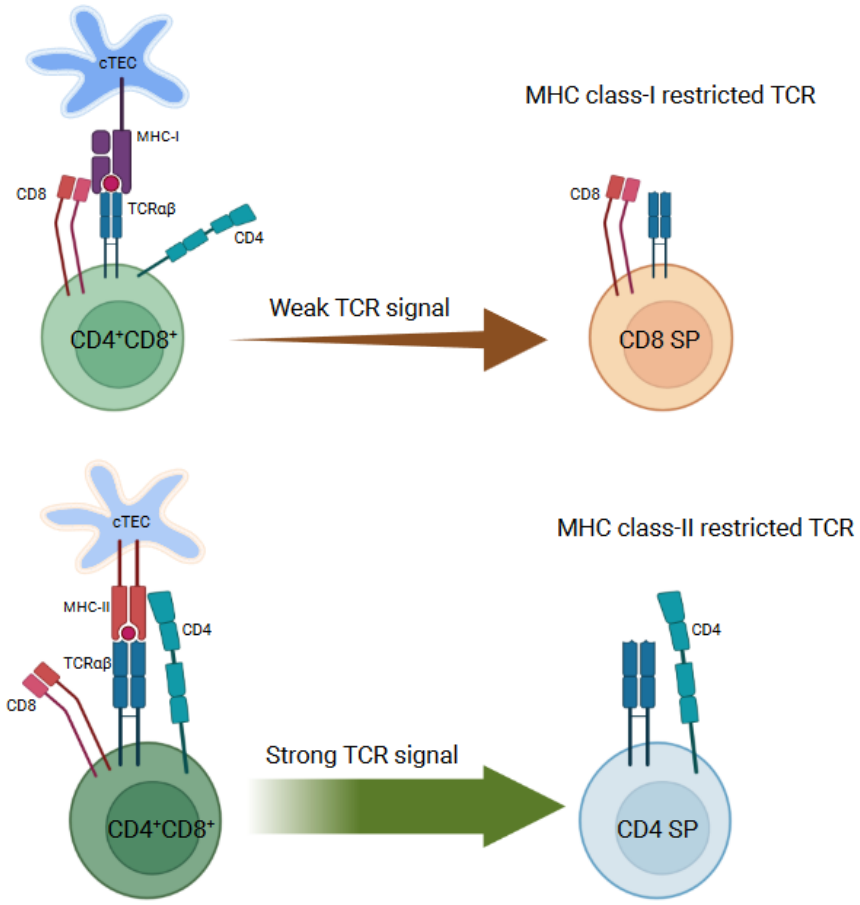


Figure 3.1.2: Strength of signal model of lineage choice. According to the strength-of-signal instructive paradigm, Cd8 expression is stopped by strong TCR signals whilst Cd4 transcription is stopped by mild TCR stimuli. Mature CD4+ T cells are produced by robust signaling from CD4 along with pMHC-class-II:TCRs, while mature CD8+ T cells are produced by weak signaling from CD8 alongside pMHC-class-I:TCRs.

The duration of signal model is based on the strength of signal hypothesis, suggesting the importance of signal duration in addition to the strength of the signal. The model postulates that extended TCR signaling directs DP thymocytes to stop the expression of Cd8 gene and $CD4^+$ SP program is initiated, whereas brief TCR signals direct DP thymocytes to stop expressing the *Cd4* gene and allows their differentiation into $CD8^+$ SP cells (Figure 3.1.3). The major premise of the duration-of-signal model is that somehow the duration/strength of TCR signal varies between TCR-CD8:MHC I and TCR-CD4:MHC II co-engagement. According to

the duration-of-signal model (Aliahmad and Kaye, 2006; Kappes et al., 2005, 2006), signaled DP thymocytes eventually emerge as $CD4^{+}CD8^{low}$ regardless of continuous production of both Cd4 and Cd8 transcripts explained by the signaling dependent variation on their cell-surface deposition (Lucas and Germain, 1996). However, experimental systems using endogenous: constitutive-transgenic alleles of Cd8 gene expression suggests that the endogenous expression of Cd8 mRNA is downregulated in all DP cells upon receiving the TCR signal (Bosselut et al., 2003). This could not be reconciled with the duration model.

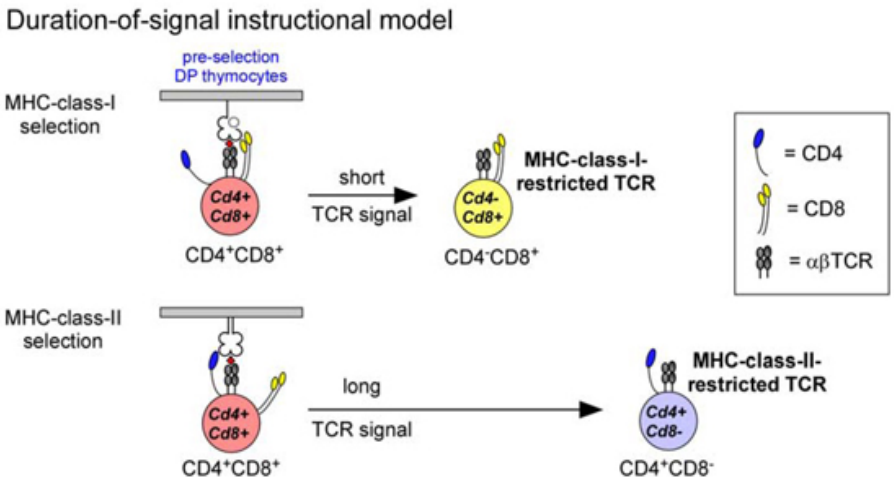


Figure 3.1.3: The duration model of CD4/CD8 lineage choice. According to the duration-of-signal instructive paradigm, Cd8 transcription is stopped by prolonged and even strong TCR signals, while Cd4 expression is stopped by brief or even low intensity TCR signals. Reproduced from Singer et al., 2008.

3.1.1.2 The kinetic signalling model: Current understanding in the CD4/CD8 lineage choice paradigm

While all the classical models discussed above have numerous experimental evidences against their central premises, the kinetic signaling hypothesis is built on the empirical data and incorporates the assistance of cytokine signaling in the overall scheme (Figure 3.1.4). The kinetic-signaling model posits that DP cells that have received TCR signals initially cease the transcription of Cd8 gene, and then they evaluate how TCR signaling is affected if Cd8 expression is reduced (Brugnera et al., 2000; Singer, 2002; Singer and Bosselut, 2004). The DP cells develop into $CD4^{+}$ T cells if the TCR-mediated positive selection cues are active in absence of Cd8 expression. Contrastingly, the thymocytes develop into $CD8^{+}$ T cells if the

TCR-mediated positive selection signaling stops due to lack of Cd8 expression. This also suggests that the positive selection signals need to occur prior to CD4/CD8 lineage choice. The kinetic signaling paradigm also suggests that the TCR signal duration is sensed by IL-7 and other γc type cytokines that in turn are regulated by the continued activity or termination of the TCR signaling (Singer, 2002; Singer and Bosselut, 2004).

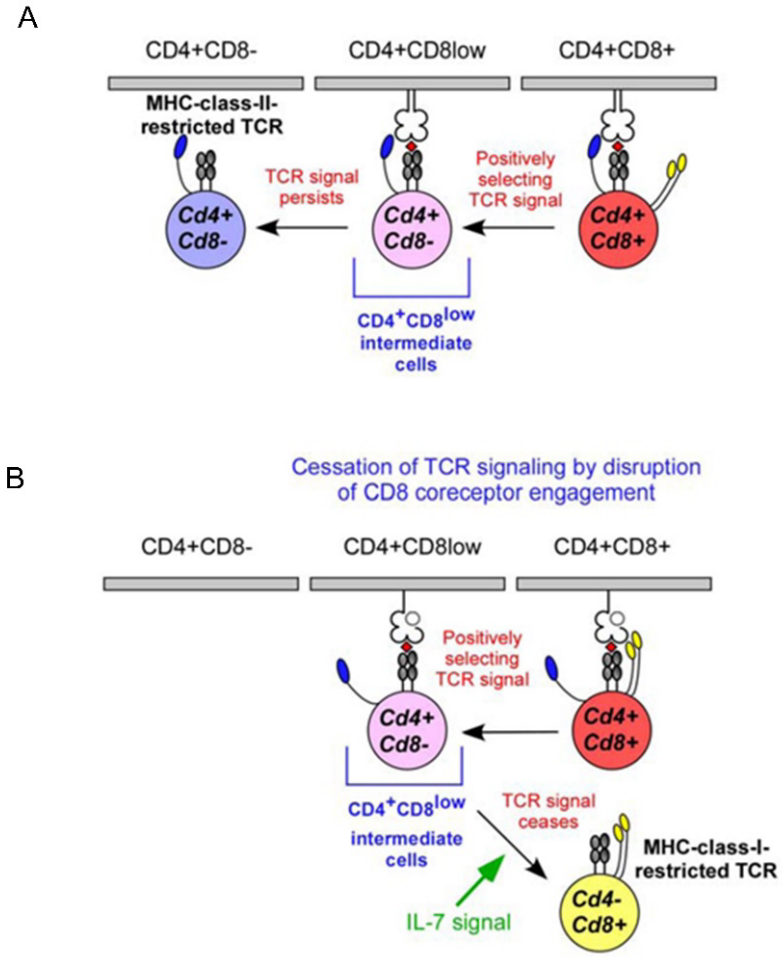


Figure 3.1.4: CD4/CD8 fate decision by kinetic signaling mechanism. A. Positively selected TCR stimuli transform the DP cells into Cd4+Cd8− intermediate cells through which lineage decision occurs. The choice depends on whether the signaling persists or not. pMHC-class-II: TCR signaling continues in Cd4+Cd8− intermediate thymocytes differentiate into $CD4^+$ T cells. **B.** Since CD8 expression is required to sustain pMHC-class-I:TCR signaling, it terminates in the $Cd4^+Cd8^-$ transitional thymocytes. IL-7 signaling promotes co-receptor reversal and these cells develop into $CD8^+$ thymocytes. Reproduced from Singer et al., 2008.

3.1.2 Factors involved in the TCR-dependent selection process via TCR activation

In the last few decades, researchers have focused on the transcription factors and chromatin regulators being important in determining the lineage fate of thymocytes (Rothenberg, 2014). T-cell specific nuclear factor for activated T cells (NFAT) is involved in regulation of various T cell specific genes such as Il2 and Il4 (Chow et al., 1999). NFAT consists of a family of at least four specific members, carries out multiple stage specific functions such as positive selection (Hogan et al., 2003; Koltsova et al., 2007) and differentiation (Hogan et al., 2003; Teixeira et al., 2005). Additionally, NFAT cooperates with transcriptional complex AP1 and NFκB on the chromatin to mediate transcription of target genes (Macian et al., 2001; Napetschnig and Wu, 2013). These and other lineage specific transcription factors bind to cis-regulatory elements as well as distal enhancers and brings about transcription (Rothenberg, 2014). As mentioned above, the regulation of many genes provides a positive-feedback to downstream genes, gene regulation at a global epigenetic level assumes a central role in thymocyte development (Lee et al., 2001). In case where DNMT1 was conditionally ablated in mouse thymocytes, impairment in TCR mediated proliferation of DP cells, with an irregular gene expression in naïve CD4 cells was observed (Lee et al., 2001). Additionally, the SWI/SNF complex, involved in nucleosome remodeling, when knocked-out in T cells resulted in dysregulation in number of CD4 or CD8 co-receptors on the thymocyte surface (Chi et al., 2002). Thus, such global chromatin modifiers are indeed necessary for normal development of T cells. The major events after TCR:APC interaction are shown in Figure 3.1.5.

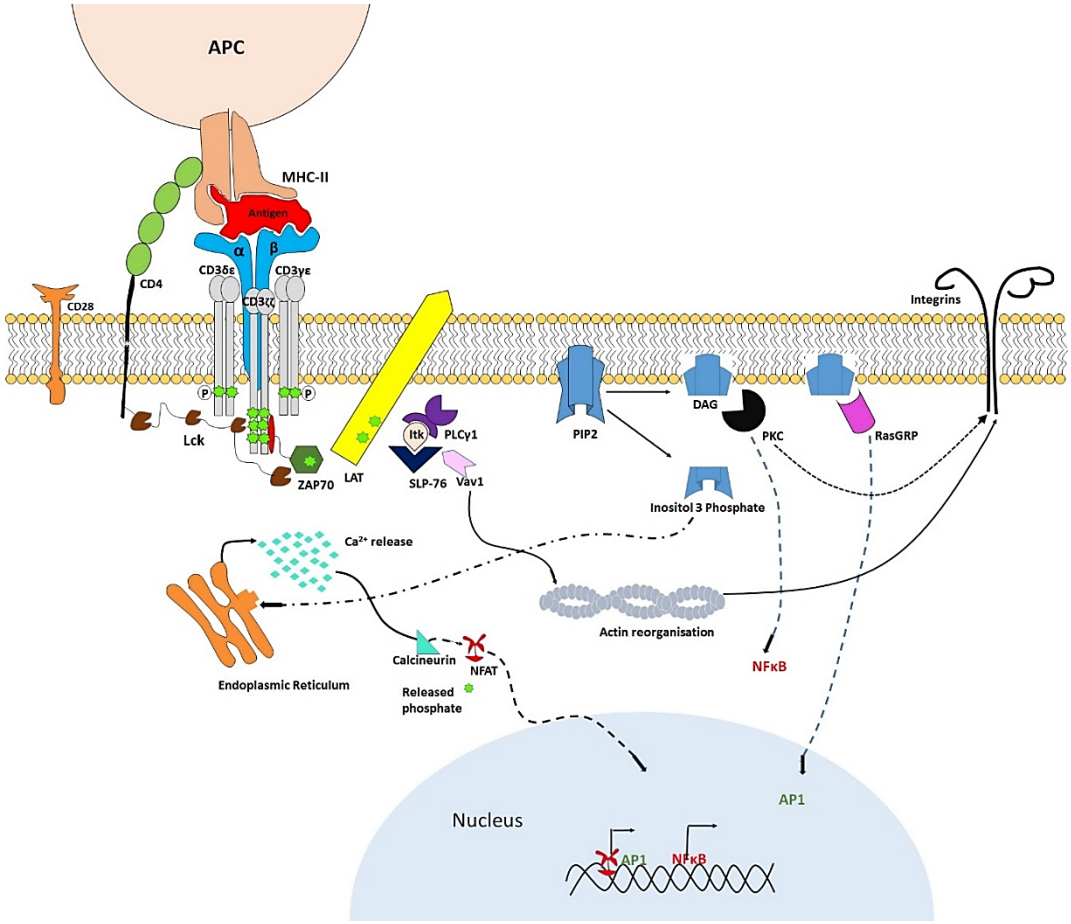


Figure 3.1.5: Simplified molecular view of T cell activation cascade upon T cell activation. Upon stimulation of TCR, two parallel pathways – NFκB and NFAT get activated via the $Ca2^{+}$ signaling and PKC mediated activation, respectively. These two pathways along with AP1 complex mediate transcription of various T cell activation and proliferation genes. Cellular structural changes are mediated by actin reorganization which helps in cell migration and allows transport of surface proteins.

One such factor is Special AT rich binding protein (SATB1) which is a T-cell enriched chromatin organizer and transcription factor (Pavan Kumar et al., 2006). SATB1 is a nuclear protein, expression of which ensues in the DN4 stage, further increasing in DP and CD4 stages (Gottimukkala et al., 2016). We have previously shown that SATB1 is positively regulated upon TCR stimulus in naïve CD4 T cells (Gottimukkala et al., 2016). It was further demonstrated that SATB1 has exhibits a 'bimodal' distribution in CD4 cells, which is directly dependent on the activation status of the CD4 cells as marked by CD69 levels. Additionally, both TCRβ and

CD69 positive cells demonstrate the highest expression of SATB1, corroborating the role of TCR signaling in regulating its expression (Gottimukkala et al., 2016). SATB1 loops the MHC class-I locus into transcriptionally active region, reducing the requirement of many transcriptionally active units (Galande et al., 2007; Kumar et al., 2007). Looping and transcriptional regulation of Th2 specific cytokine genes has also been shown (Cai et al., 2006). SATB1 has been shown to interacts with β-catenin promoting differentiation of Th2 CD4 lymphocytes (Notani et al., 2010). β-catenin is an important mediator of the Wnt pathway, regulating an array of genes required in proliferation, survival and development (MacDonald et al., 2009). Thus, SATB1 is known to be an important link between TCR signaling and Wnt pathway (Notani et al., 2010). SATB1 has also been implicated in lymphoid lineage commitment of HSCs (Satoh et al., 2013).

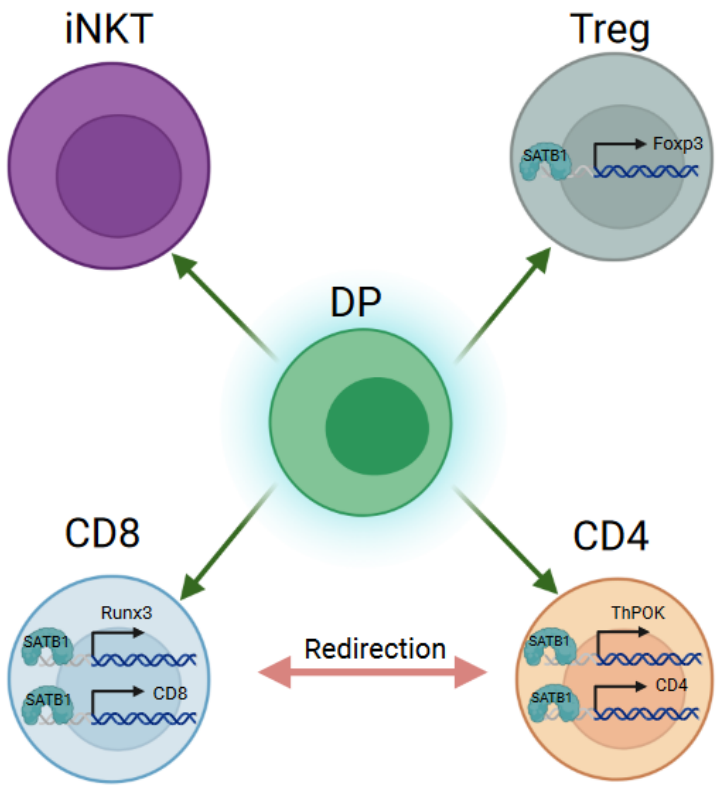


Figure 3.1.6: SATB1 is essential for multiple aspects of T cell development. Investigations utilizing SATB1 conditional knockout showed a developmental arrest at the double positive (DP) stage, which further decreased the quantity of differentiating T cells. The key TFs during thymocyte development after DP stage mentioned are bound by SATB1, by which their expression is regulated. Adapted from Kakugawa et al., 2017.

Using conditional knockout (cKO) of SATB1, it has recently been shown that positive selection is partly affected by SATB1 deficiency (Kondo et al., 2016). Also, SATB1 is shown to be essential for ThPOK expression (Kakugawa et al., 2017) as well as the de-repression of CD8a gene (Figure 3.1.6). Furthermore, it has recently been demonstrated that SATB1 regulates PD1 expression by directly binding to its promoter region (Stephen et al., 2017). Despite all these studies, the molecular mechanisms of SATB1's role in the development of thymocytes as well as naïve T cell activation has not been fully understood.

In this chapter, we aimed to investigate the effect of TCR signaling on SATB1 mediated gene expression and its cooperativity with other signaling induced transcription factors. We show that SATB1 is an early TCR gene, expression of which is directly proportional to the strength of the TCR signal. We further show that upon TCR signaling, SATB1 interacts with NFAT1 and NFAT2. NFAT and SATB1 are further demonstrated to bind multiple TCR-dependent genes, and regulate their expression in a cooperative manner.

3.2 Materials and methods

3.2.1 Mice

Six- to ten-week-old WT C57/B6 mice were used to prepare the single cell suspension of thymocytes for the flow cytometry analysis as well as for sorting the subpopulations of thymocytes which are in different stages of T-cell development. All the strains were bred and maintained in sterile environment and the dissections were carried out in accordance with the guidelines of NFGFHD, IISER Pune.

3.2.2 Immunostaining

A single-cell suspension of thymocytes was made using thymi isolated from six weeks old C57/b6 mice. In order to fix thymocytes, 2% paraformaldehyde was used. Permeabilization of fixed thymocytes was performed with 0.1% Triton X-100. Following permeabilization, thymocytes were treated for three hours at room temperature with mouse anti-SATB1 (Abcam) and rabbit anti-NFAT1 (Abcam). Thymocytes were stained intracellularly and then washed with 1X PBS containing 0.01% Tween 20. The secondary antibodies were stained for one hour at room temperature using the fluorochrome-tagged secondary antibodies, anti-mouse GFP 488 and anti-rabbit AlexaFluor 594. For intracellular staining of CD4 T cells, AF 488 tagged Nur77 (eBioscience) and AF-647 SATB1 (BD Biosciences) were used as primary antibodies, and a separate secondary staining and DAPI staining steps were omitted. Following the staining with a secondary antibody, DNA was labeled with DAPI (Sigma). Using diluted cell suspension, slides were prepared using Cytospin (Thermo Scientific). Images were acquired using Anisotropy microscopy (Zeiss).

3.2.3 Western blotting

BCA method was used to quantify the proteins after the cells were lysed using RIPA lysis buffer (50 mM Tris-HCl at PH 7.4, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, 0.25% sodium deoxycholate, 1 mM PMSF, and 1X protease inhibitor cocktail (Roche). After the whole protein was separated on a 10 %SDS polyacrylamide gel, it was transferred to a Millipore PVDF membrane. Blocking of non-specific interactions was done with 5% skimmed milk. the following antibodies were used to probe the blot: anti-Actin (1:1000, Millipore), anti-SATB1 (1:1000, Santacruz Biotechnologies), and anti-NFAT1(1:1000) at 4°C overnight. Next day, multiple washes were given to the blots with 1x TBST (500mM NaCl, 20mM Tris pH 7.5, 1mM EDTA, 1% Tween 20) and anti-primary secondary antibodies tagged with HRP (anti-mouse

IgG-HRP or anti-rabbit-HRP) were used to probe the blots followed by extensive washing with 1X TBST. Using the ImageQuant LAS 4000 system (GE Healthcare Life Sciences) and the ECL luminescence detection reagent (BIO-RAD), the signals were shown.

3.2.4 Cell culture and transfection

Growth media supplemented with 10% fetal bovine serum (Gibco BRL, Grand Island, NY, USA) was used to cultivate HEK 293T cells, which were then kept in a humidified incubator at 37°C and 5% CO2. Transfection of cloned NFAT1 and SATB1, SATB1 and NFAT2 was performed using Lipofectamine 3000 (Invitrogen). Cells were cultured for 48 hrs. Harvested cells were washed twice with cold 1x PBS and the pellets were either resuspended in 1 mL Trizol (Invitrogen) followed by snap freezing in liquid nitrogen or directly stored at -80°C.

3.2.5 Isolation of naïve $CD4^+$ T cells and cell culture

A single-cell suspension of the spleen was made, and RBCs were lysed using RBC lysis solution in order to isolate naïve $CD4^+$ T-cells. Using a mouse naïve $CD4^+$ T cell isolation kit (BD Biosciences), cells were sorted magnetically. Briefly, the antibody cocktail was added to the total cell suspension and incubated on ice for 30 min. Then the magnetic particles were added to the antibody-bound cell suspension and incubated at 4°C for 60 min. The bound cells were discarded and the flow through was collected as pure CD4 population.

A gradient of .125, 0.25, 0.5, 1 and 2 µg/mL of plate-bound anti-mouse CD3 (Clone 17A2, eBioscience) and 1.5 µg/mL of soluble anti-mouse CD28 (Clone 37.51, eBioscience) were added to naïve $CD4^+$ T-cell cultures for a duration of 48 h. The cells were harvested and subjected to either western blotting or RT-qPCR.

3.2.6 Flow cytometry

The suspension of thymocytes and or splenocytes were prepared using thymii from 6-week-old mice from various genotypes indicated and were used for surface staining. Before staining, Fc -blocking was done with anti-CD16.32 antibody (BD Biosciences). Then, the T cells were subjected to surface staining using the following the fluorochrome tagged antibodies: BV 421 anti-mouse CD4 (Biolegend); PE anti-mouse CD8a (BD Biosciences); followed by internal staining with FITC Nur77 (eBioscience) and AF-647 SATB1 (BD Biosciences). The flow cytometry analysis was performed using Celesta analyzer (BD Biosciences).

3.2.7 Nucleofection and cell culture

To isolate the thymus, C57BL/6 mice from 6 weeks old were used. A single cell suspension was made by crushing the thymus through a 70-micron filter. Ten million thymocytes were electroporated with 400 nM of scr or si SATB1 RNA using Lonza IIb system (Lonza) with the preset 'Mouse CD4'. The transfected thymocytes were grown in RPMI-1640 medium supplemented with 10% FBS, penicillin-streptomycin, and IL7 at 5ng/mL, SCF at 5 ng/mL, and IL2 at 2 ng/mL for 24 h. Following incubation, cells were collected and subjected to either western blotting or RNA isolation.

3.2.8 Chromatin Immunoprecipitation (ChIP) q RT-PCR

Thymocytes from control, Satb1 knockdown and NFAT inhibited samples were used for chromatin immunoprecipitation (ChIP). DP thymocytes from WT and Satb1 cKO mice were used for chromatin immunoprecipitation (ChIP). For ten minutes, 50×10^6 sorted DP thymocytes were cross-linked at room temperature using 1% formaldehyde which was quenched using 125 mM glycine. Buffer 1 (25 mM Tris-HCl pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.1% NP40, 1 mM DTT, 0.5 mM PMSF and 1X protease inhibitor cocktail (PIC) (Roche) was used to isolate the nuclei from the cross-linked cells. The nuclei were resuspended in lysis buffer (50 mM Tris-HCl pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X–100, 0.1% Sodium deoxycholate, 0.1% SDS, 0.5 mM PMSF, and 1X PIC). Sonication of the chromatin was performed to get an average fragment distribution of 200–600 base pairs. The chromatin was then subjected to sonication using a Covaris M220 ultrasound sonication system for 12 min, peak power 75 and burst cycle of 200. Protein A/G magnetic bead slurry (Thermo) was used to preclear the sonicated chromatin for 1 h. The beads were discarded and the precleared chromatin was subjected to antibody binding using anti-mouse SMC1A (Abcam) antibody overnight at 4°C. The antibody-bound protein (and associated chromatin) was precipitated by adding Protein A/G beads (saturated using 10 mg/mL tRNA and 1% BSA and at 4°C for 4h). The lysate was discarded and the beads were washed thoroughly using the lysis buffer. Elution of complexed from A/G beads was done using the elution buffer (1% SDS, 0.1 M NaHCO3). The eluted chromatin was de-crosslinked and the treated with the proteinase K and RNase A. Further the DNA was purified and subjected to quantitative RT-PCR (qRT-PCR) analysis using the Sybr green qRT-PCR master mix (Takara) at the following PCR conditions: 95°C for 5 min, 95°C, 60°C, 72°C- repeated for 40 cycles. The ChIP occupancies were normalized with 0.1% Input DNA using the ΔΔCT method. The following primers were used for qRT-PCRs:

Gene	Sequence
BATF F	TGGTCTTTTCCCAGATGTGTAT
BATF R	GTCCTCCTCTCTCTCGGGT
mIL2Ra F	TCACAGAACAGAGTAGGCACA
mIL2Ra R	CTCAGGCCTCTCAGTCTGTC
Dusp22 F	CCCCTTGACGTCTGCCAC
Dusp22 R	AAGCCATGTCAGTTCTAGGGT
ETS F	TTCAAGGACCTTTGGTACGG
ETS R	TTCTAGGTTGAGGGGGAGGT

3.2.9 Co-Immunoprecipitation Immunoblotting

Thymocyte single-cell suspension was made from thymii isolated from six-week-old mice. Thermo Scientific's BCA kit was used to assess the protein content after cells were lysed using NP40-based lysis buffer. Using 3 µg mouse or rabbit IgG and protein A/G Dyna beads (Thermo Scientific), 500 µg of protein was precleared for 1 hr. Using anti-SATB1 antibody, the cleared lysate was subjected to antibody-binding for four hours at 4°C. Following incubation, protein A/G magnetic beads (Thermo Scientific) were used to precipitate anti-SATB1-bound protein complexes. At least 3 washes were given to the beads with lysis buffer. 1% SDS in 1X PBS was used to elute proteins at 98°C for 15 min. Co-Immunoprecipitated proteins were identified by subjecting the eluted protein to SDS-polyacrylamide gel electrophoresis followed by western blotting using SMC1A, FLAG and NFAT1 antibodies.

3.2.10 cDNA synthesis and quantitative PCR analysis (qRT-PCR)

Isolation of total RNA from thymocytes or sorted CD4 T cells subjected to various treatments was performed using the Trizol method. Briefly, the samples stored in Trizol reagent (Themo Scientific) were thawed at RT, followed by addition of chloroform. The samples were mixed thoroughly by vortexing and centrifuged. The top aqueous layer was collected and equal volume of Isopropanol was added. The sample was mixed gently and incubated at RT for 10 min. Centrifugation was done and supernatant was discarded. The pellets were washed with chilled 70% ethanol. The pellets were dried at RT and resuspended in nuclease free water. After DNase I (Roche) treatment, the RNA was cDNA synthesis was done using High-capacity cDNA synthesis kit (Applied Biosystems). Quantitative PCR was carried out with qPCR master mix (Takara) using the below PCR conditions: 95°C-5 min, 95°C-30 sec, 60°C-45 sec, 72°C-1 min for 40 cycles. Normalization was done using 18s rRNA levels as an internal control. The following qPCR primers were used:

Gene	Sequence
18s rRNA F	TGTGCCATGGTGTCAAGTGT
18s rRNA R	TAGCAAGGCTGACCAAGGTG
IRF4 F	CTCATGTGGAACCTCTGCTAG
IRF4 R	GTGGTAATCTGGAGTGGTAACG
DUSP22 F	ACAGTTGAGCAGGAACAAGG
DUSP22 R	CTTGTCAGGTTTTGAGATGGTG
IL2 F	GGAACCTGAAACTCCCCAGG
IL2 F	AATCCAGAACATGCCGCAGA
Ctla4 F	CACGTCTCCAGGTCCTCAG
Ctla4 R	GAGAAAGGAAGCCGTGGGT
Bcl2 F	CCTGTGGATGACTGAGTACCTG
Bcl2 R	AGCCAGGAGAAATCAAACAGAGG
SATB1	GTTTTCTGCGTGGTGGAACA
SATB1	GCAGAGCTGTGTGAATAGCC

3.3 Results

3.3.1 Satb1 expression is dependent on the strength of TCR signal

It has been previously shown that SATB1 expression is governed positively by TCR signaling (Patta and Madhok et al., 2020). We thus probed to see whether thymocytes, under activating conditions mimicking TCR activation i.e. via phorbol ester PMA and $Ca2^{+}$ ionophore ionomycin addition (Smith-Garvin et al., 2009), show variation in expression of SATB1. We observed that SATB1 levels indeed increased upon activation of both Jurkat T cell line as well as in developing T cells (Fig 3.3.1A and Fig 3.3.1B). Calcium signaling alone via the ionophore addition was sufficient for SATB1 induction in T cells (Fig 3.3.1B). Furthermore, we analyzed publicly available RNA-seq data of CD4 T cells in naïve and anti-CD3 activated conditions. We observed that the levels of Satb1 increases dramatically upon CD4 T cell activation, which is confirmed by Nur77 expression (Fig 3.3.1C)

It has been shown that T cell specific transcription factor BATF is under direct control of TCR signal strength (Iwata et al., 2017), upon which it recruits its transcription partner IRF4 to mediate differential transcriptional regulation (Iwata et al., 2017). It has previously been postulated that TCR strength along with signal kinetics might play a role in developing T cells (Iwata et al., 2017; Singer et al., 2008). We therefore checked whether Satb1 is governed by the strength of TCR signal and whether it can regulate gene expression differentially under such conditions. We utilized an in vitro approach of T cell activation using anti-CD3 and anti-CD28 treatment. We found that upon activation of Naïve CD4 T cells, there is a dramatic increase in SATB1 levels at the transcript (Fig 3.3.1D) and the protein (Fig 3.3.1E) levels. We observe that SATB1 levels exhibit a direct influence by TCR signal strength, wherein with increasing TCR signaling SATB1 is upregulated at both transcript and protein levels (Fig 3.3.1D and 3.3.1E). Further, SATB1 levels saturate at a specific concentration and decrease beyond it presumably due to activation induced apoptotic response of CD4 cells. Furthermore, flow cytometry analysis upon in vitro activation of Naïve CD4 T cells corroborated the increase in protein levels based on TCR signal strength (Fig 3.3.1F). Activation status of these cells was quantified by NR4A (Nurr77) expression (Fig 3.3.1G), which is a known TCR strength marker (Ashouri and Weiss, 2017). These results demonstrate that SATB1 expression is dependent on TCR signal strength and might regulate gene expression depending on the surface signal.

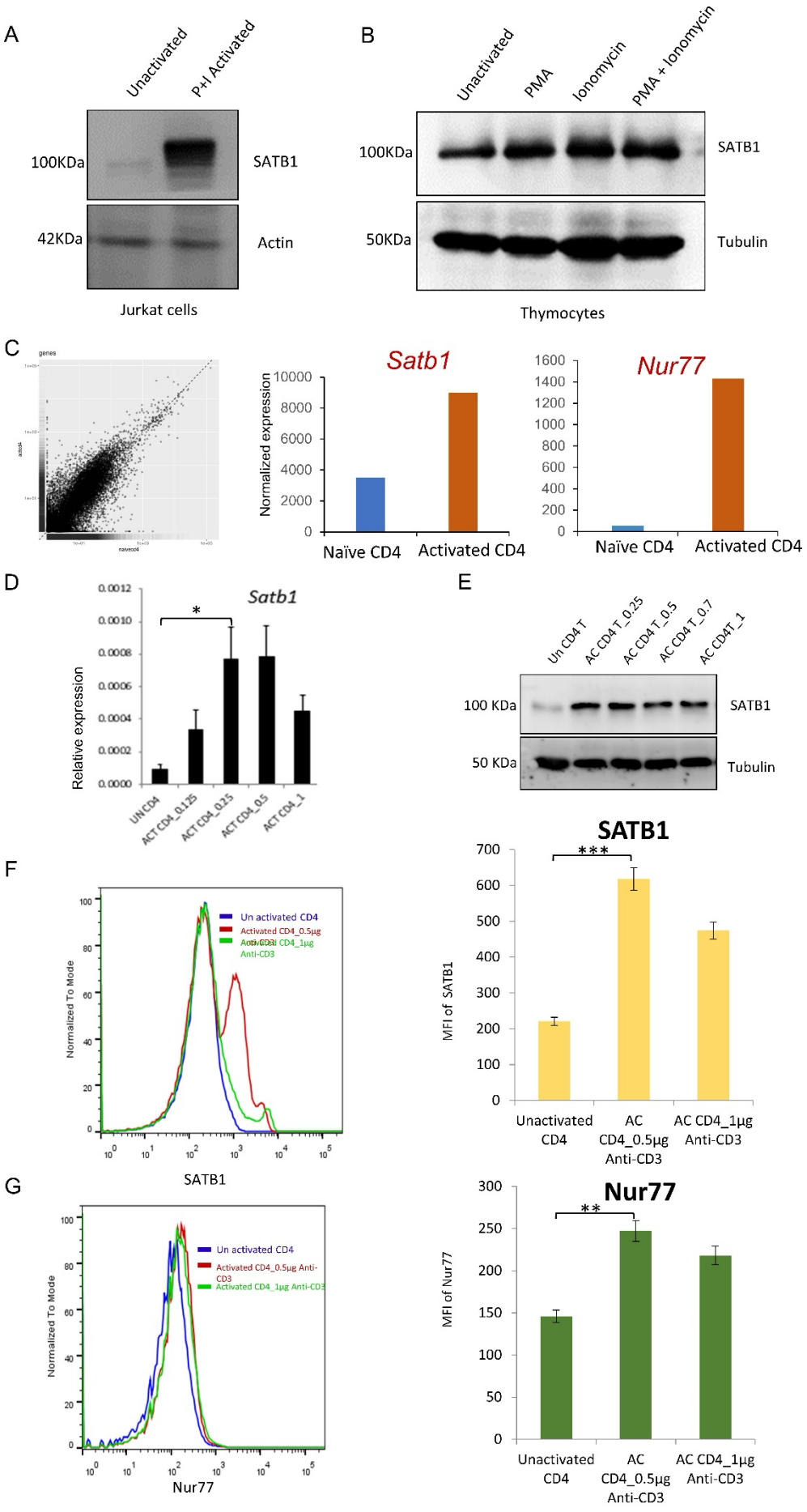

A
Unactivated
P+I Activated
100KDa
SATB1
42KDa
Actin
Jurkat cells
B
Unactivated
PMA
Ionomycin
PMA + Ionomycin
100KDa
SATB1
50KDa
Tubulin
Thymocytes
C
genes
Satb1
Normalized expression
Naïve CD4
Activated CD4
Nur77
Naïve CD4
Activated CD4
D
Satb1
Relative expression
UN CD4
ACT CD4_0.125
ACT CD4_0.25
ACT CD4_0.5
ACT CD4_1
E
Un CD4 T
AC CD4 T_0.25
AC CD4 T_0.5
AC CD4 T_0.7
AC CD4T_1
100 KDa
SATB1
50 KDa
Tubulin
F
Un activated CD4
Activated CD4_0.5µg Anti-CD3
Activated CD4_1µg Anti-CD3
Normalized To Mode
SATB1
SATB1
MFI of SATB1
Unactivated CD4
AC CD4_0.5µg Anti-CD3
AC CD4_1µg Anti-CD3
G
Un activated CD4
Activated CD4_0.5µg Anti-CD3
Activated CD4_1µg Anti-CD3
Normalized To Mode
Nur77
Nur77
MFI of Nur77
Unactivated CD4
AC CD4_0.5µg Anti-CD3
AC CD4_1µg Anti-CD3

Figure 3.3.1: TCR signal strength dictates SATB1 expression. A. Jurkat T cell line was treated with PMA and Ionomycin. Western blotting shows an increase in SATB1 expression upon stimulation of T cells. **B.** Freshly isolated thymocytes were stimulated with PMA, Ionomycin or PMA and Ionomycin. Immunoblotting shows the sensitivity of SATB1 expression to individual as well as combinatorial activation in vitro. **C.** RNA seq analysis of mature CD4 T cells in naïve versus activated conditions (GSE34550). Expression of SATB1 and activation marker Nur77 is shown. **D.** CD4 T cells were isolated and cultured with different concentrations of anti-CD3 as shown, with anti-CD28 as constant. The RT-qPCR analysis shows the dependency of SATB1 on the amounts of anti-CD3 thus, TCR signal strength. **E.** Western blotting analysis of SATB1 upon treatment with anti-CD3 gradient described in **D**, corroborating the transcript level data. **F.** Internal staining of SATB1 was done on cultured naïve and activated CD4 T cells. The flow cytometry analysis shows a TCR signal dependent increase in SATB1 protein. **G.** The activation status of the treated T cells was confirmed with Nur77 internal staining and flow cytometry analysis.

3.3.2 SATB1 activates TCR induced genes in a strength-dependent manner

To assess the control of TCR signal strength on SATB1 target genes, we used the same linear gradient of anti-CD3, keeping anti-CD28 constant. We analyzed publicly available Satb1 ChIP-seq data in developing T cells (GSE64407). A few direct targets of SATB1 were selected–Bcl2 and Ctla4, as shown by Satb1 binding on their promoters (Fig 3.3.2A). Further assessment using qRT-PCR showed that the transcript levels of these genes also show TCR signal strength dependency (Fig 3.3.2B). Interestingly Ctla4 transcripts plateaus at the saturation concentration of TCR strength and is dependent on Satb1 only for TCR strength dependent induction.

To delineate the functional significance of Satb1 on TCR dependent genes, we knocked down (KD) Satb1 using custom RNAi. We monitored the efficiency of KD in HEK293T cells using si-RNA and sh-RNA constructs (Fig 3.3.2C). Further, we nucleofected thymocytes with control and Satb1 si RNA. We confirmed the KD of Satb1 in thymocytes at ~60% efficiency by western blotting (Fig 3.3.2D). We used a portion of the KD-validated samples to perform qRT-PCR of Satb1 bound target genes. We found that specific TCR induced genes, which have a direct binding site for Satb1 on their promoters, are positively regulated by Satb1 upon TCR signaling such as Il2, Dusp22 and Irf4 (Fig 3.3.2E). These results suggest that Satb1 directly targets

TCR induced genes and regulates their expression in a TCR signal strength-dependent manner.

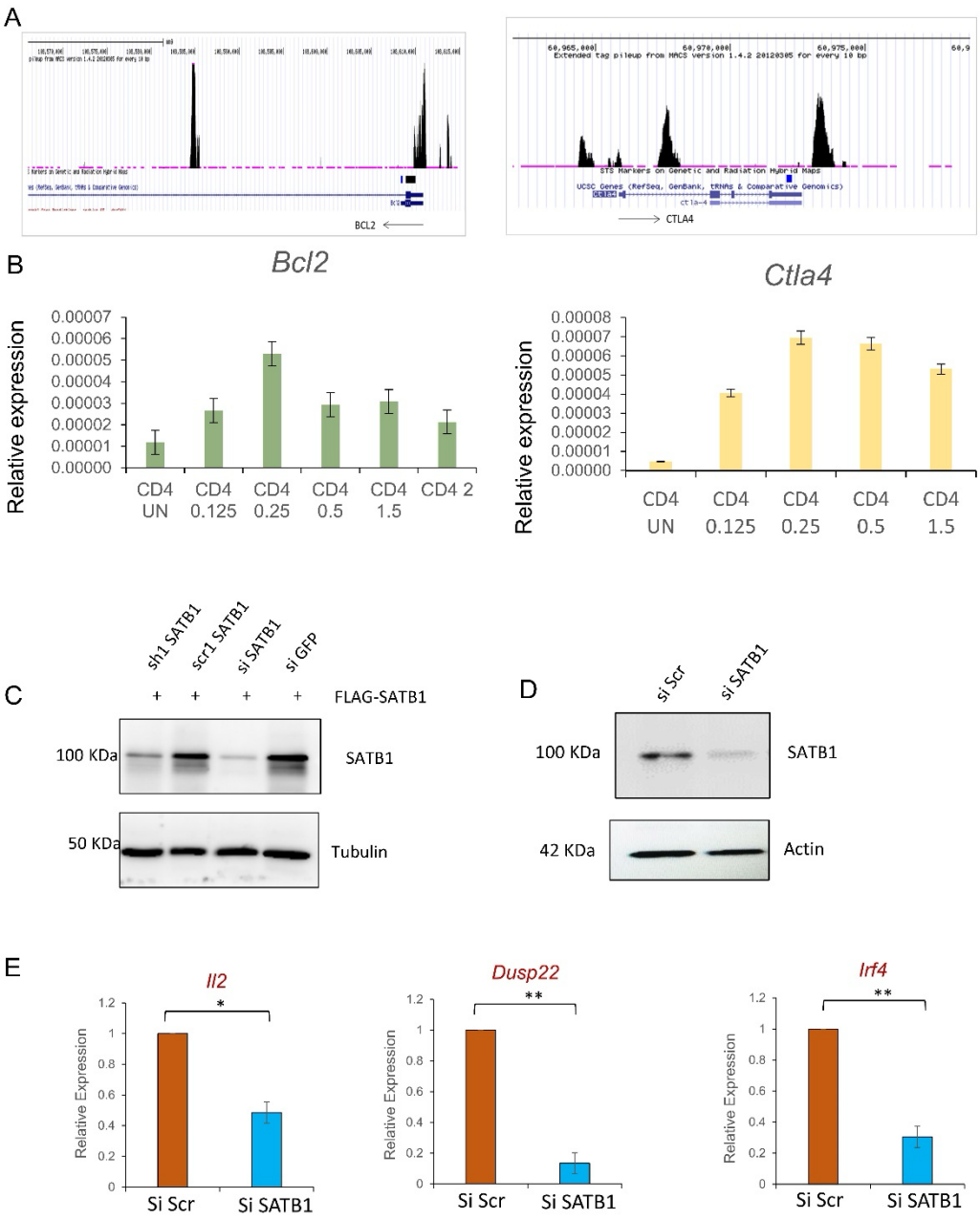


Figure 3.3.2: SATB1 activates TCR induced genes in a strength-dependent manner. A. ChIP-seq of SATB1 in developing T cells was analyzed (GSE64407). ChIP-seq occupancy profiles of SATB1 on *Bcl2* (Left) and *Ctla4* (right) promoters are shown. **B.** qRT-PCR analysis for Bcl2 and Ctla4 expression in naïve and treated (as indicated) CD4 T cells showing a signal strength dependency of SATB1 direct targets. **C.** Validation of custom synthesized siRNA for SATB1 knockdown (KD) and compared with Satb1 shRNA mediated knockdown in HEK293T. **D.** siRNA mediated knockdown was performed by nucleofection of thymocytes. The KD

efficiency was upto 50%. **E.** Expression of a few SATB1 targets (identified from ChIP-seq) upon SATB1 depletion is shown. n=3. Graphpad v10 was used for the statistical analysis, unpaired Student's t-Test was performed * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

3.3.3 Satb1 interacts with NFAT1 and NFAT2 in developing T cells

SATB1 is shown to interact with numerous proteins in T cells including PML (Kumar et al., 2007), CtBP1 (Purbey et al., 2009) and β-catenin (Notani et al., 2010), via its N-terminal domain, with each interaction leading to activation/repression of different set of genes. Thus, we hypothesized that SATB1 under TCR signal might interact with T cell specific transcription factors to facilitate activation of these cells as a 'fine tuner' (Fu et al., 2014), and thus would help in effective selection of T cells in thymic microenvironment.

Upon activation of T cells, in response to $Ca2^{+}$ Nuclear factor of Activated T cells (NFAT) diffuses to nucleus and kick starts the activation of various cytokine genes involved in survival and proliferation (Macian, 2005). In parallel, AP1 complex protein FOS, along with NFκB family proteins translocate to the nucleus upon receiving TCR stimulus via Ras and PKC pathways respectively, and mediate transcription co-operatively with NFAT family proteins (Smith-Garvin et al., 2009). This general scheme of transcriptional activation is conserved in thymocytes as well, wherein they receive variable TCR signaling during as their development progresses in the thymus. We thus probed to check if SATB1 interacts with TCR specific responder proteins, which specifically translocate to the nucleus and begin numerous transcription events. NFAT1 is known to be involved in T cell differentiation and development with some level of redundancy with other NFAT family members (NFAT2 NFAT3 and NFAT4) (Macian, 2005). Thus, we first performed immunostaining of SATB1 in total thymocytes along with NFAT1, under resting (Fig 3.3.3A) and PMA-Ionomycin (P+I) stimulated conditions (Fig 3.3.3B). We observed that Satb1 and Nfat1 co-localize in the nucleus of developing T cells. Furthermore, we confirmed the physical interaction between Nfat1 and Satb1 by co-immunoprecipitation in both conditions (Fig 3.3.3C and Fig 3.3.3D). As shown, P+I induced activation dramatically increased the colocalization as well as the pull-down efficiency, suggesting the interaction may be activation dependent. We also cloned Satb1, NFAT1 and NFAT2 in different mammalian expression vectors. Upon overexpression of Satb1 with either NFAT1 or NFAT2 in HEK293T, and performing co-immunoprecipitation, we showed that SATB1 interacts with both NFAT1 and NFAT2 (Fig 3.3.3 E and Fig 3.3.3F). These findings indicate that SATB1 physically interacts with TCR-induced transcription factors NFAT1/2 in a signal-dependent manner.

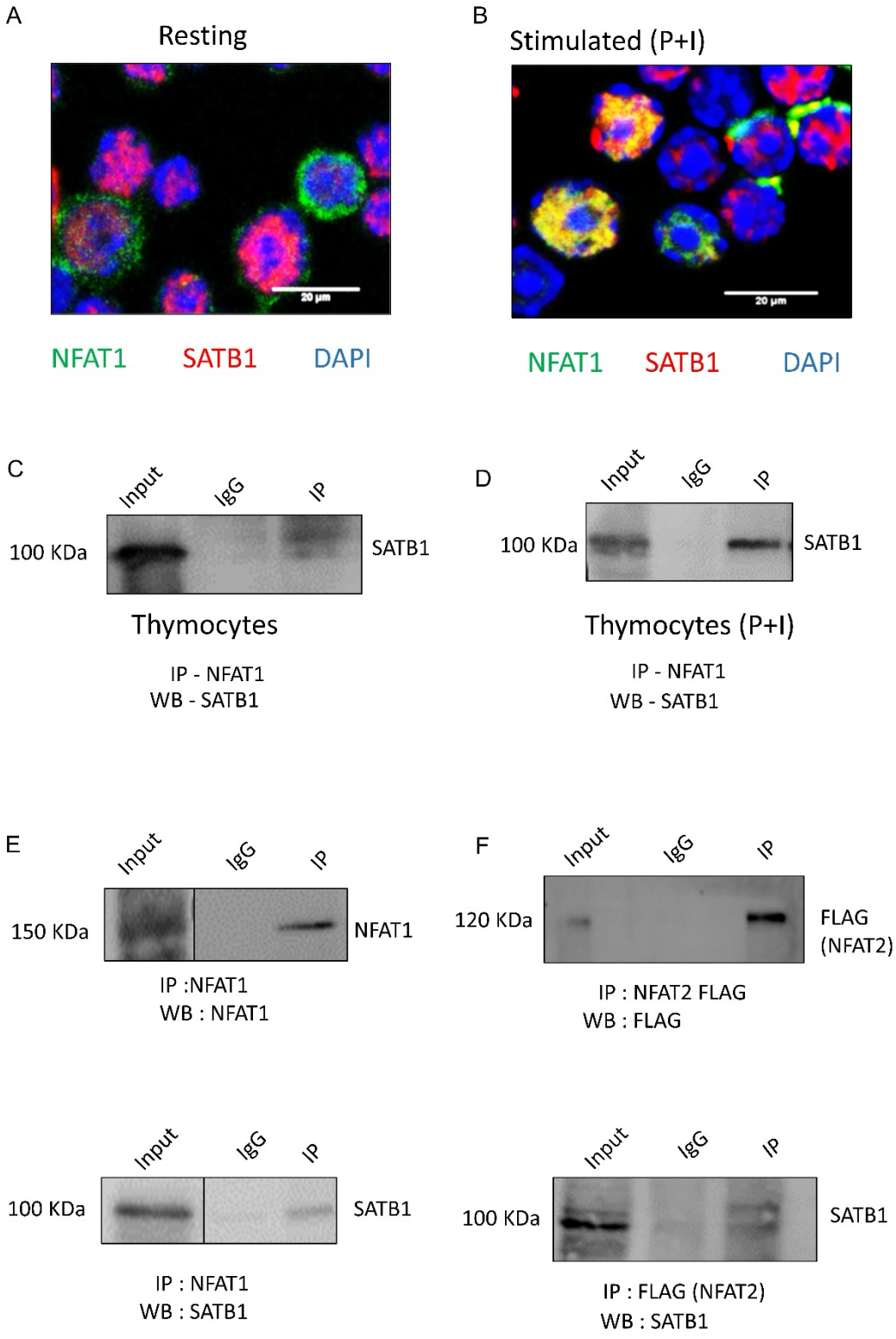


Figure 3.3.3: Satb1 interacts with NFAT1 and NFAT2 in developing T cells. A and **B.** Resting and PMA+Ionomycin (P+I) stimulated thymocytes were stained internally with SATB1 (red), NFAT1 (green) and DAPI (blue) and were imaged using confocal microscope. The increase in colocalization of NFAT and SATB1 upon P+I stimulation is represented. **C.** Co-

immunoprecipitation (co-IP) of SATB1 along with NFAT1 in resting thymocytes, followed by immunoblotting (IB) of SATB1. **D.** Similarly, co-IP of SATB1 with NFAT1 was performed under P+I stimulated condition and the interaction is shown via SATB1 IB. The increase in the intensity of interaction is in part due to NFAT translocation to the nucleus upon P+I stimulation. **E.** NFAT1 and SATB1 were ectopically expressed in HEK293T. Co-IP of NFAT1 and SATB1 is shown further confirming their physical interaction. **F.** Similar to **E**, 3X Flag-tagged NFAT2 and SATB1 were overexpressed in HEK293T and co-IP of SATB1 with FLAG was performed confirming that SATB1 interacts with more than one isoform of NFAT. For both **E** and **F**, top blots demonstrate the efficiency of IPs and bottom blots depict the co-immunoprecipitated protein (n=3).

3.3.4 Satb1 cooperates with NFAT to facilitate the expression of TCR induced genes

Transcription factors are known to interact on the chromatin (Macian, 2005; Macian et al., 2001; Notani et al., 2010), and could recruit various histone modifiers and RNA polymerase, as well as different regulatory elements (Hsu et al., 2015). Recently, it has been shown that transcription factor PU.1 is important in recruiting other lineage specific transcription factors to sites other than their canonical binding sites, thus mediating differential gene regulation (Hosokawa et al., 2018).

SATB1 is known to act in coordination with the Wnt signaling mediator β-catenin to drive differentiation of CD4 T cells to the helper-2 T lineage (Th2) (Notani et al., 2010), whereas knockdown (KD) of SATB1 results dysregulation of Wnt β-catenin target genes. As a chromatin organizer, SATB1 is involved in the dynamic organization of numerous genes. It has been shown that knockout of SATB1 leads to downregulation of RAG1, resulting in inefficient VDJ recombination (Hao et al., 2015). Coordinated gene regulation at specific loci via SATB1 mediated looping is reported (Cai et al., 2006; Kumar et al., 2007; Yasui et al., 2002). NFAT, on the other hand, is primarily studied for regulation of key cytokines such as IL-2 –NFAT having binding sites on several critical regions of Il2 promoter (Spolski et al., 2018). NFAT also cooperates with AP1 complex proteins cFOS and cJUN, as well as NFkB to begin active gene expression (Smith-Garvin et al., 2009).

To assess the cooperation ability of SATB1 and NFAT on gene regulation, we performed ChIP-qPCRs for SATB1 for some of the common gene targets of NFAT and SATB1 (Fig 3.3.4A). Similarly, ChIP-qPCRs for NFAT1 were also performed on the promoters of the same target

genes (Fig 3.3.4B). We find that of the many common target genes tested, there is high enrichment of both SATB1 and NFAT1 on *Ets1* and *Dusp22* promoters. Further, we used the same experimentation in NFAT inhibited (NFATi) and SATB1 KD conditions i.e. we performed SATB1 ChIP for NFATi treated thymocytes, and NFAT1 ChIP for SATB1 KD samples. We observed a dramatic reduction in the occupancy of SATB1 on the promoter of Dusp22 upon NFAT inhibition (Fig 3.3.4C). Similarly, SATB1 KD resulted in decreased occupancy of NFAT1 on the Dusp22 promoter. Thus, suggesting that NFAT1 is required for SATB1 recruitment on the promoters of their commonly bound targets and vice versa.

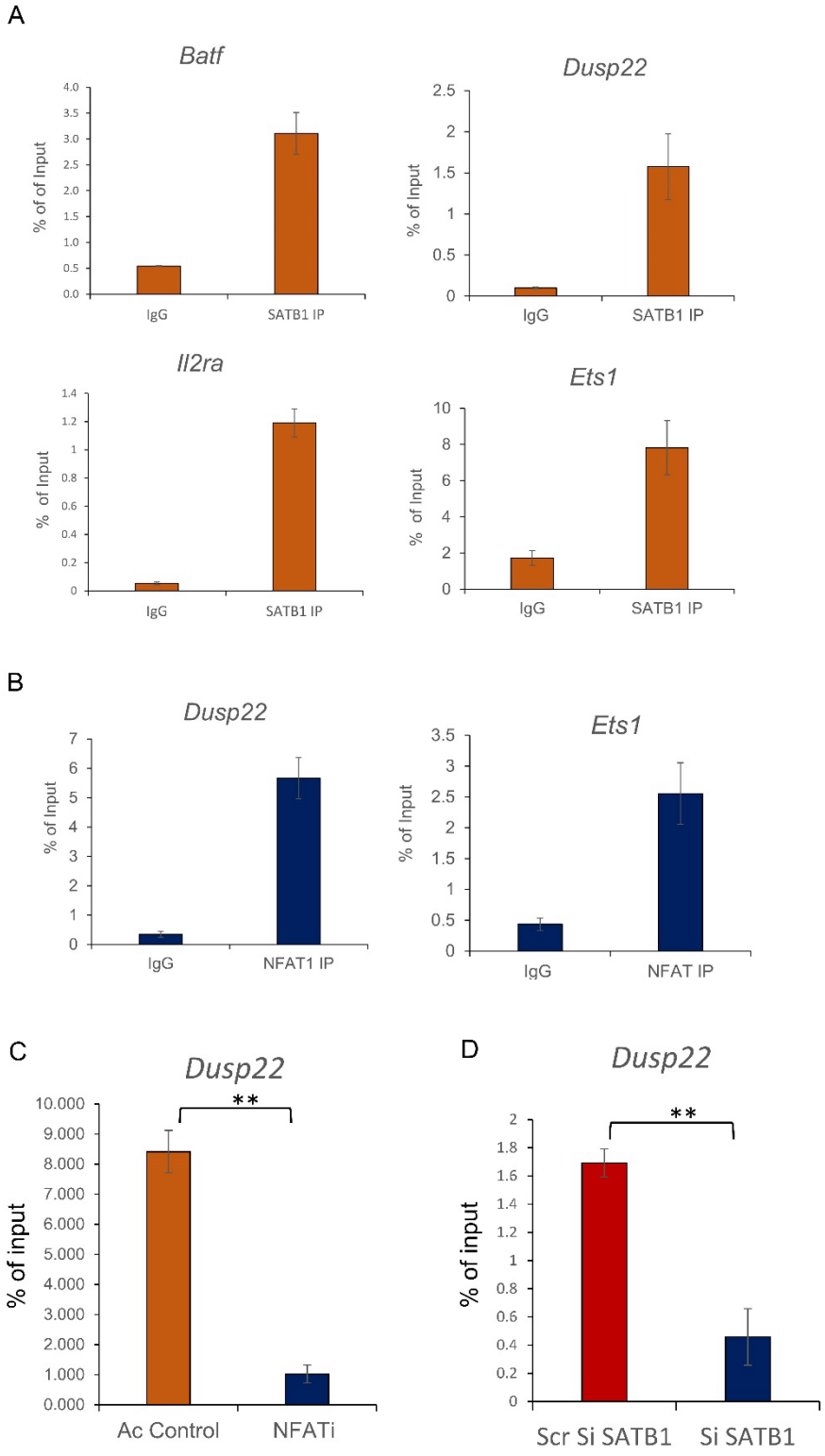


Figure 3.3.4: Satb1 cooperates with NFAT to facilitate the expression of TCR induced genes. A. ChIP-qPCR analysis on some of the TCR induced genes is shown, confirming the enrichment of SATB1 on the promoters of these genes. **B.** ChIP-qPCR analysis of NFAT1 scoring for the same genomic regions as analyzed in A, showing NFAT occupancy on the same binding region as SATB1. **C.** NFAT nuclear translocation was inhibited with a specific

inhibitor -N7032. ChIP RT-qPCR was performed for SATB1. The difference in chromatin occupancy of SATB1 at *Dusp22* promoter is shown in control and NFATi conditions. **D.** Similarly, SATB1 knockdown was performed in thymocytes, followed by NFAT1 ChIP. RT-qPCR analysis shows a decreased occupancy of *NFAT1* on *Dusp22* promoters upon SATB1 depletion (n=3). Graphpad v10 was used for the statistical analysis, unpaired Student's T-test was performed * $p<0.05$, ** $p<0.01$.

3.4 Discussion

A key determinant of T cell responses upon TCR:pMHC interaction -both in terms of magnitude and kinetics - is the strength of the TCR signal. T cell development during positive selection is, at least partly, dependent on the TCR affinity differences in vivo (Zeidan et al., 2019). SATB1 is a critical regulator of T cell development as well as a lineage defining factor for Th2 differentiation (Notani et al., 2010). It has been shown that SATB1 expression is upregulated upon TCR signal induction (Gottimukkala et al., 2016).

Our results show that SATB1 expression is sensitive to the strength of TCR signal in vivo and in vitro. Naïve CD4 T cells shut of SATB1 expression as the mature and migrate from the thymus to the secondary lymphoid organs. Our results demonstrate that SATB1 is linearly sensitive to the TCR signal strength in vitro both at transcript and protein levels. At higher activation levels, ie. beyond saturation of receptor ligation, SATB1 expression decreases in coherence with the reduction in T cell activation. Thus, SATB1 expression is highly sensitive to the changes in the ‘degree’ of T cell activation. SATB1 direct targets also show a similar trend of TCR strength dependence in vitro.

With the activation of TCR signaling CD4 T cells begin to express SATB1, likely via a promoter switch, as shown by our previous data (Patta et al., 2020). SATB1 harbors four distinct promoters - P1 through P4. We earlier demonstrated that specifically upon TCR signal induction there is a switch in SATB1 promoter usage in CD4 T cells-specifically shifts from P1 to P2 expression, which in turn leads to more efficient translation of SATB1 protein (Patta et al., 2020). The TCR-induced switch in selective transcription of P2 over P1 is mediated by TCF1, which directly binds to P2 at multiple sites and positively regulates its expression. Differential activation by varying TCR strength might regulate SATB1 promoter usage differently, which could in turn impart the dosage sensitivity of SATB1 in mediating transcriptional control.

In SATB1 null condition, the T lymphocyte development was severely impaired, specifically, stagnated at the DP stage (Alvarez et al., 2000). This was further confirmed by conditionally knocking out SATB1 from HCSs and thus, subsequent lineages (Kondo et al., 2016; Satoh et al., 2013). Although, many genes important for T cell development such as Rag1/2, which were affected by SATB1 KO in HSCs, were restored as development proceeded (Satoh et al.,

2013), suggesting compensating mechanisms. We show that upon knockdown of SATB1 results in downregulation of various TCR-dependent genes in developing T cells such as Batf, Ets1, Il2 and Ctla4, suggesting that SATB1 may act as an early regulator of TCR signaling. The varying affinity of TCR:pMHC interactions during positive selection as well as during naïve CD4/CD8 activation could quantitatively affect the expression of direct as well as indirect targets of SATB1 in its absence.

SATB1, being a high affinity matrix attachment region (MAR)-binding protein, is mostly localized in the nucleus and remains bound to chromatin. It associates with a multitude of factors with diverse, even opposing functions in terms of gene activation in a signal- dependent manner. For instance, SATB1 is shown to interact with HDAC1 (Kumar et al., 2005; Pavan Kumar et al., 2006) in resting CD4 T cells maintaining the basal transcription. Whereas, upon TCR activation SATB1 association with HDAC is replaced by P300 and PCAF leading to de-repression of SATB1 bound genes (Pavan Kumar et al., 2006). We show that upon the activation of TCR signaling in vitro, SATB1 associates with NFAT1 and NFAT2 proteins. Although the interaction was not observed in our IP-MS data (discussed in Chapter 1), probably due to very low expression of NFAT family proteins in thymocytes. As the $Ca2^{+}$ signaling is initiated downstream of the TCR, NFAT gets translocated to the nucleus (Macian, 2005). We demonstrate that SATB1 associates with NFAT1 on the chromatin and co-bind various TCR 'inducible' genes. Upon further probing into their cooperative association on the target loci, we show that SATB1 is necessary for NFAT1 to bind a few of these targets such as Dusp22. Similarly, upon inhibiting NFAT translocation to the nucleus by inhibiting $Ca2^{+}$ signaling, we demonstrated that NFAT binding to *Dusp22* promoter is essential for SATB1 occupancy.

In conclusion, we demonstrate that TCR signal strength positively dictates SATB1 expression in a dose-dependent manner, which in turn regulate the direct targets of SATB1 in a similar way. We further identify SATB1 as a positive regulator to TCR signaling as knockdown of SATB1 results in reduction of TCR activated genes. We demonstrate that SATB1 physically interacts with nuclear factor of activated T cells (NFAT) family proteins NFAT1 and NFAT2. Finally, we show a synergy between NFAT and SATB1 in the activation of Dusp22 gene via direct co-dependent promoter binding. Interestingly, NFAT expression is also positively regulated by TCR signaling although it is not sensitive to TCR signal strength (Gallagher et al., 2021). Thus, SATB1 tuning via TCR strength could act as a 'rheostat' to their co-binding to common targets and mediate a signal strength-dependent gene regulation.

Satb1 was first characterized as a chromatin binding protein with a strong affinity for base unpairing regions of DNA (Dickinson et al., 1997; Dickinson et al., 1992; Galande et al., 2001). Satb1 has been shown to interact with multiple factors (Cai et al., 2003; Cai et al., 2006; Yasui et al., 2002), although functional relevance of such interactions was studied in the past two decades by multiple groups including prominent contributions from our group (Kumar et al., 2007; Mir et al., 2016; Notani et al., 2010; Pavan Kumar et al., 2006; Purbey et al., 2009). As discussed in previous chapters, Satb1 exhibits an unusual protein structure, with both highly-structured and intrinsically disordered interaction arms. A locus specific looping activity of Satb1 has been reported by multiple groups (Cai et al., 2006; Hao et al., 2015; Kumar et al., 2007). Recent studies have shown that Satb1 has a 3D genome organizing effect, albeit to a minor extent (Feng et al., 2022; Zelenka et al., 2022). Our findings suggest that Satb1 interacts with transcriptional activators and cohesin, functioning at both shorter promoter-enhancer and long-range looping levels. As demonstrated by Galande group (Notani et al., 2010) and others (Hosokawa et al., 2018), Satb1's dynamic binding motif allows it to bind broader DNA elements compared to other transcription factors. It is possible that Satb1 acts as a 'rheostat' for cohesin complex binding, as its absence affects the loading and eviction of Smc1a at its targets and non-targets, respectively. The DNA binding of Satb1 and its proximity to other factors might influence the extent of its interactions.

During T cell activation, increased Satb1 expression, along with other factors, plays a crucial role in regulating the transcriptional turnover of immune-specific genes. Given the importance of the Calcineurin-NFAT pathway in T cell activation, regulating its efficacy and target specificity is essential. We hypothesize that upon T cell activation, NFAT enters the nucleus and interacts with Satb1, which co-binds to its DNA targets. Although we haven’t provided genome-wide binding data, our target-specific binding assessment suggests a cooperative nature between Satb1 and NFAT on shared targets in activated T cells. Satb1 induction is an early event in T cell activation, with protein levels increasing within 2 hours of PMA and Ionomycin treatment, possibly via the Erk pathway. Previous research from Galande laboratory has shown that post-translational modifications of Satb1 shift from phosphorylation to acetylation (Pavan Kumar et al., 2006; Purbey et al., 2009), altering its interactome.

Based on the expression status of the three complexes, we propose that Satb1's association with cohesin may change due to shifts in 3D chromatin architecture. Its binding to TCR-induced genes may increase due to NFAT interaction and acetylation. Acetylation of Satb1 might promote its association with NFAT, but this remains to be tested. We believe that increased Satb1 expression upon activation induces the phase transition property of the

protein via IDRs. During phase separation, Satb1 may incorporate local transcription factors like NFAT and the cFos-cJun complex at the promoters of its target genes, facilitating effective transcription. Our data suggest that not all Satb1 target genes are affected by its phase separation property, possibly due to its foci formation ability near specific targets where Pol II hubs are enriched upon activation.

In summary, Satb1 interaction with the cohesin complex maintains a steady state in T cells. Upon activation, chromatin rewiring increases Satb1's interaction with NFAT, leading to the binding of this complex at TCR-induced gene promoters and transcription. Interaction of the cohesin complex with Satb1 persists and may be necessary to induce key TCR-specific genes, such as Cd3e, through chromatin accessibility changes requiring their association.

3.5 References

Aliahmad, P., and Kaye, J. (2006). Commitment issues: linking positive selection signals and lineage diversification in the thymus. Immunol Rev *209*, 253-273.

Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev *14*, 521-535.

Ashouri, J.F., and Weiss, A. (2017). Endogenous Nur77 Is a Specific Indicator of Antigen Receptor Signaling in Human T and B Cells. J Immunol *198*, 657-668.

Bell, J.J., and Bhandoola, A. (2008). The earliest thymic progenitors for T cells possess myeloid lineage potential. Nature *452*, 764-767.

Bosselut, R., Feigenbaum, L., Sharrow, S.O., and Singer, A. (2001). Strength of signaling by CD4 and CD8 coreceptor tails determines the number but not the lineage direction of positively selected thymocytes. Immunity *14*, 483-494.

Bosselut, R., Guinter, T.I., Sharrow, S.O., and Singer, A. (2003). Unraveling a revealing paradox: Why major histocompatibility complex I-signaled thymocytes "paradoxically" appear as CD4+8lo transitional cells during positive selection of CD8+ T cells. J Exp Med *197*, 1709-1719.

Brugnera, E., Bhandoola, A., Cibotti, R., Yu, Q., Guinter, T.I., Yamashita, Y., Sharrow, S.O., and Singer, A. (2000). Coreceptor reversal in the thymus: signaled CD4+8+ thymocytes

initially terminate CD8 transcription even when differentiating into CD8+ T cells. Immunity *13*, 59-71.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat Genet *38*, 1278-1288.

Carpenter, A.C., and Bosselut, R. (2010). Decision checkpoints in the thymus. Nat Immunol *11*, 666-673.

Chan, S.H., Cosgrove, D., Waltzinger, C., Benoist, C., and Mathis, D. (1993). Another view of the selective model of thymocyte selection. Cell *73*, 225-236.

Chan, S.H., Waltzinger, C., Baron, A., Benoist, C., and Mathis, D. (1994). Role of coreceptors in positive selection and lineage commitment. EMBO J *13*, 4482-4489.

Chi, T.H., Wan, M., Zhao, K., Taniuchi, I., Chen, L., Littman, D.R., and Crabtree, G.R. (2002). Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. Nature *418*, 195-199.

Chow, C.W., Rincon, M., and Davis, R.J. (1999). Requirement for transcription factor NFAT in interleukin-2 expression. Mol Cell Biol *19*, 2300-2307.

Davis, C.B., Killeen, N., Crooks, M.E., Raulet, D., and Littman, D.R. (1993). Evidence for a stochastic mechanism in the differentiation of mature subsets of T lymphocytes. Cell *73*, 237-247.

Erman, B., Alag, A.S., Dahle, O., van Laethem, F., Sarafova, S.D., Guinter, T.I., Sharrow, S.O., Grinberg, A., Love, P.E., and Singer, A. (2006). Coreceptor signal strength regulates positive selection but does not determine CD4/CD8 lineage choice in a physiologic in vivo model. J Immunol *177*, 6613-6625.

Fu, G., Rybakin, V., Brzostek, J., Paster, W., Acuto, O., and Gascoigne, N.R. (2014). Fine-tuning T cell receptor signaling to control T cell development. Trends Immunol *35*, 311-318.

Galande, S., Purbey, P.K., Notani, D., and Kumar, P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr Opin Genet Dev *17*, 408-414.

Gallagher, M.P., Conley, J.M., Vangala, P., Garber, M., Reboldi, A., and Berg, L.J. (2021). Hierarchy of signaling thresholds downstream of the T cell receptor and the Tec kinase ITK. Proc Natl Acad Sci U S A *118*.

Gascoigne, N.R., Rybakin, V., Acuto, O., and Brzostek, J. (2016). TCR Signal Strength and T Cell Development. Annu Rev Cell Dev Biol *32*, 327-348.

Germain, R.N. (2002). T-cell development and the CD4-CD8 lineage decision. Nat Rev Immunol *2*, 309-322.

Gottimukkala, K.P., Jangid, R., Patta, I., Sultana, D.A., Sharma, A., Misra-Sen, J., and Galande, S. (2016). Regulation of SATB1 during thymocyte development by TCR signaling. Mol Immunol *77*, 34-43.

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

Hogan, P.G., Chen, L., Nardone, J., and Rao, A. (2003). Transcriptional regulation by calcium, calcineurin, and NFAT. Genes Dev *17*, 2205-2232.

Holst, J., Wang, H., Eder, K.D., Workman, C.J., Boyd, K.L., Baquet, Z., Singh, H., Forbes, K., Chruscinski, A., Smeyne, R.*, et al.* (2008). Scalable signaling mediated by T cell antigen receptor-CD3 ITAMs ensures effective negative selection and prevents autoimmunity. Nat Immunol *9*, 658-666.

Hosokawa, H., Ungerback, J., Wang, X., Matsumoto, M., Nakayama, K.I., Cohen, S.M., Tanaka, T., and Rothenberg, E.V. (2018). Transcription Factor PU.1 Represses and Activates Gene Expression in Early T Cells by Redirecting Partner Transcription Factor Binding. Immunity *49*, 782.

Hsu, H.T., Chen, H.M., Yang, Z., Wang, J., Lee, N.K., Burger, A., Zaret, K., Liu, T., Levine, E., and Mango, S.E. (2015). TRANSCRIPTION. Recruitment of RNA polymerase II by the pioneer transcription factor PHA-4. Science *348*, 1372-1376.

Itano, A., Kioussis, D., and Robey, E. (1994). Stochastic component to development of class I major histocompatibility complex-specific T cells. Proc Natl Acad Sci U S A *91*, 220-224.

Itano, A., and Robey, E. (2000). Highly efficient selection of CD4 and CD8 lineage thymocytes supports an instructive model of lineage commitment. Immunity *12*, 383-389.

Itano, A., Salmon, P., Kioussis, D., Tolaini, M., Corbella, P., and Robey, E. (1996). The cytoplasmic domain of CD4 promotes the development of CD4 lineage T cells. J Exp Med *183*, 731-741.

Iwata, A., Durai, V., Tussiwand, R., Briseno, C.G., Wu, X., Grajales-Reyes, G.E., Egawa, T., Murphy, T.L., and Murphy, K.M. (2017). Quality of TCR signaling determined by differential affinities of enhancers for the composite BATF-IRF4 transcription factor complex. Nat Immunol *18*, 563-572.

Janeway, C.A., Jr. (1992). The T cell receptor as a multicomponent signalling machine: CD4/CD8 coreceptors and CD45 in T cell activation. Annu Rev Immunol *10*, 645-674.

Kakugawa, K., Kojo, S., Tanaka, H., Seo, W., Endo, T.A., Kitagawa, Y., Muroi, S., Tenno, M., Yasmin, N., Kohwi, Y.*, et al.* (2017). Essential Roles of SATB1 in Specifying T Lymphocyte Subsets. Cell Rep *19*, 1176-1188.

Kappes, D.J., He, X., and He, X. (2005). CD4-CD8 lineage commitment: an inside view. Nat Immunol *6*, 761-766.

Kappes, D.J., He, X., and He, X. (2006). Role of the transcription factor Th-POK in CD4:CD8 lineage commitment. Immunol Rev *209*, 237-252.

Kappler, J.W., Roehm, N., and Marrack, P. (1987). T cell tolerance by clonal elimination in the thymus. Cell *49*, 273-280.

Koltsova, E.K., Ciofani, M., Benezra, R., Miyazaki, T., Clipstone, N., Zuniga-Pflucker, J.C., and Wiest, D.L. (2007). Early growth response 1 and NF-ATc1 act in concert to promote thymocyte development beyond the beta-selection checkpoint. J Immunol *179*, 4694-4703.

Kondo, M., Tanaka, Y., Kuwabara, T., Naito, T., Kohwi-Shigematsu, T., and Watanabe, A. (2016). SATB1 Plays a Critical Role in Establishment of Immune Tolerance. J Immunol *196*, 563-572.

Krangel, M.S. (2009). Mechanics of T cell receptor gene rearrangement. Curr Opin Immunol *21*, 133-139.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A., and Galande, S. (2007). Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat Cell Biol *9*, 45-56.

Kumar, P.P., Purbey, P.K., Ravi, D.S., Mitra, D., and Galande, S. (2005). Displacement of SATB1-bound histone deacetylase 1 corepressor by the human immunodeficiency virus type 1 transactivator induces expression of interleukin-2 and its receptor in T cells. Mol Cell Biol *25*, 1620-1633.

Lee, P.P., Fitzpatrick, D.R., Beard, C., Jessup, H.K., Lehar, S., Makar, K.W., Perez-Melgosa, M., Sweetser, M.T., Schlissel, M.S., Nguyen, S.*, et al.* (2001). A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. Immunity *15*, 763-774.

Leung, R.K., Thomson, K., Gallimore, A., Jones, E., Van den Broek, M., Sierro, S., Alsheikhly, A.R., McMichael, A., and Rahemtulla, A. (2001). Deletion of the CD4 silencer element supports a stochastic mechanism of thymocyte lineage commitment. Nat Immunol *2*, 1167-1173.

Love, P.E., Lee, J., and Shores, E.W. (2000). Critical relationship between TCR signaling potential and TCR affinity during thymocyte selection. J Immunol *165*, 3080-3087.

Lucas, B., and Germain, R.N. (1996). Unexpectedly complex regulation of CD4/CD8 coreceptor expression supports a revised model for CD4+CD8+ thymocyte differentiation. Immunity *5*, 461-477.

MacDonald, B.T., Tamai, K., and He, X. (2009). Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell *17*, 9-26.

Macian, F. (2005). NFAT proteins: key regulators of T-cell development and function. Nat Rev Immunol *5*, 472-484.

Macian, F., Lopez-Rodriguez, C., and Rao, A. (2001). Partners in transcription: NFAT and AP-1. Oncogene *20*, 2476-2489.

Masuda, K., Kakugawa, K., Nakayama, T., Minato, N., Katsura, Y., and Kawamoto, H. (2007). T cell lineage determination precedes the initiation of TCR beta gene rearrangement. J Immunol *179*, 3699-3706.

Murphy, K.M., Heimberger, A.B., and Loh, D.Y. (1990). Induction by antigen of intrathymic apoptosis of CD4+CD8+TCRlo thymocytes in vivo. Science *250*, 1720-1723.

Napetschnig, J., and Wu, H. (2013). Molecular basis of NF-kappaB signaling. Annu Rev Biophys *42*, 443-468.

Notani, D., Gottimukkala, K.P., Jayani, R.S., Limaye, A.S., Damle, M.V., Mehta, S., Purbey, P.K., Joseph, J., and Galande, S. (2010). Global regulator SATB1 recruits beta-catenin and regulates T(H)2 differentiation in Wnt-dependent manner. PLoS Biol *8*, e1000296.

Palmer, E. (2003). Negative selection--clearing out the bad apples from the T-cell repertoire. Nat Rev Immunol *3*, 383-391.

Patta, I., Madhok, A., Khare, S., Gottimukkala, K.P., Verma, A., Giri, S., Dandewad, V., Seshadri, V., Lal, G., Misra-Sen, J.*, et al.* (2020). Dynamic regulation of chromatin organizer

SATB1 via TCR-induced alternative promoter switch during T-cell development. Nucleic Acids Res *48*, 5873-5890.

Pavan Kumar, P., Purbey, P.K., Sinha, C.K., Notani, D., Limaye, A., Jayani, R.S., and Galande, S. (2006). Phosphorylation of SATB1, a global gene regulator, acts as a molecular switch regulating its transcriptional activity in vivo. Mol Cell *22*, 231-243.

Purbey, P.K., Singh, S., Notani, D., Kumar, P.P., Limaye, A.S., and Galande, S. (2009). Acetylation-dependent interaction of SATB1 and CtBP1 mediates transcriptional repression by SATB1. Mol Cell Biol *29*, 1321-1337.

Robey, E., Itano, A., Fanslow, W.C., and Fowlkes, B.J. (1994). Constitutive CD8 expression allows inefficient maturation of CD4+ helper T cells in class II major histocompatibility complex mutant mice. J Exp Med *179*, 1997-2004.

Rothenberg, E.V. (2014). Transcriptional control of early T and B cell developmental choices. Annu Rev Immunol *32*, 283-321.

Satoh, Y., Yokota, T., Sudo, T., Kondo, M., Lai, A., Kincade, P.W., Kouro, T., Iida, R., Kokame, K., Miyata, T.*, et al.* (2013). The Satb1 protein directs hematopoietic stem cell differentiation toward lymphoid lineages. Immunity *38*, 1105-1115.

Shaw, A.S., Amrein, K.E., Hammond, C., Stern, D.F., Sefton, B.M., and Rose, J.K. (1989). The lck tyrosine protein kinase interacts with the cytoplasmic tail of the CD4 glycoprotein through its unique amino-terminal domain. Cell *59*, 627-636.

Singer, A. (2002). New perspectives on a developmental dilemma: the kinetic signaling model and the importance of signal duration for the CD4/CD8 lineage decision. Curr Opin Immunol *14*, 207-215.

Singer, A., Adoro, S., and Park, J.H. (2008). Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. Nat Rev Immunol *8*, 788-801.

Singer, A., and Bosselut, R. (2004). CD4/CD8 coreceptors in thymocyte development, selection, and lineage commitment: analysis of the CD4/CD8 lineage decision. Adv Immunol *83*, 91-131.

Smith-Garvin, J.E., Koretzky, G.A., and Jordan, M.S. (2009). T cell activation. Annu Rev Immunol *27*, 591-619.

Spolski, R., Li, P., and Leonard, W.J. (2018). Biology and regulation of IL-2: from molecular mechanisms to human therapy. Nat Rev Immunol *18*, 648-659.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003). Positive and negative selection of T cells. Annu Rev Immunol *21*, 139-176.

Stephen, T.L., Payne, K.K., Chaurio, R.A., Allegrezza, M.J., Zhu, H., Perez-Sanz, J., Perales-Puchalt, A., Nguyen, J.M., Vara-Ailor, A.E., Eruslanov, E.B.*, et al.* (2017). SATB1 Expression Governs Epigenetic Repression of PD-1 in Tumor-Reactive T Cells. Immunity *46*, 51-64.

Swat, W., Ignatowicz, L., von Boehmer, H., and Kisielow, P. (1991). Clonal deletion of immature CD4+8+ thymocytes in suspension culture by extrathymic antigen-presenting cells. Nature *351*, 150-153.

Teixeira, L.K., Fonseca, B.P., Vieira-de-Abreu, A., Barboza, B.A., Robbs, B.K., Bozza, P.T., and Viola, J.P. (2005). IFN-gamma production by CD8+ T cells depends on NFAT1 transcription factor and regulates Th differentiation. J Immunol *175*, 5931-5939.

Vasquez, N.J., Kaye, J., and Hedrick, S.M. (1992). In vivo and in vitro clonal deletion of double-positive thymocytes. J Exp Med *175*, 1307-1316.

Wada, H., Masuda, K., Satoh, R., Kakugawa, K., Ikawa, T., Katsura, Y., and Kawamoto, H. (2008). Adult T-cell progenitors retain myeloid potential. Nature *452*, 768-772.

Wiest, D.L., Yuan, L., Jefferson, J., Benveniste, P., Tsokos, M., Klausner, R.D., Glimcher, L.H., Samelson, L.E., and Singer, A. (1993). Regulation of T cell receptor expression in immature CD4+CD8+ thymocytes by p56lck tyrosine kinase: basis for differential signaling by CD4 and CD8 in immature thymocytes expressing both coreceptor molecules. J Exp Med *178*, 1701-1712.

Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P., and Kohwi-Shigematsu, T. (2002). SATB1 targets chromatin remodelling to regulate genes over long distances. Nature *419*, 641-645.

Yui, M.A., Feng, N., and Rothenberg, E.V. (2010). Fine-scale staging of T cell lineage commitment in adult mouse thymus. J Immunol *185*, 284-293.

Zeidan, N., Damen, H., Roy, D.C., and Dave, V.P. (2019). Critical Role for TCR Signal Strength and MHC Specificity in ThPOK-Induced CD4 Helper Lineage Choice. J Immunol *202*, 3211-3225.

Chapter 4: Dynamic modulation in the transcriptome of γδ T cells upon T cell activation and cross-talk with Notch signaling

4.1 Introduction

γδ T cells are a small subgroup of T cells, accounting for about 1-5% of peripheral T cell populations (Zhao et al., 2018). γδ T cells express different TCR receptors - γ and δ chains - unlike the classical αβ T cells (Hayday et al., 1985). In further contrast, γδ T cells show aspects of both innate and adaptive immune responses (Kadivar et al., 2016), and are considered to bridge the two host-defense mechanisms (Holtmeier and Kabelitz, 2005). γδ T cells characteristically are different from their αβ counterparts not only in their TCR usage, but also in their tissue localization and MHC-independent antigen recognition (Figure 4.1.1) (Chiplunkar et al., 2009; Hayday, 2000). The predominantly found γδ T cells in circulating blood are of Vγ9Vδ2 subtype which have been widely studied in responses against microbial pathogens and cancer in humans (Fowler and Bodman-Smith, 2015). Furthermore, Vγ9Vδ2 T cells show dramatic plasticity, antigen-presentation and abundant inflammatory-cytokine production (Dar et al., 2014). In light of these features, Vγ9Vδ2 T cells are being studied extensively for cancer immunotherapy.

Human γδ T cells get activated and proliferate in response to nonpeptidic compounds derived from pathogenic microbes including *Mycobacterium tuberculosis* (Constant et al., 1994; Lang et al., 1995; Tanaka et al., 1994). γδ T cells also recognize non-peptide phosphoantigens produced via mevalonate pathway such as isopentenyl pyrophosphate (IPP) (Morita et al., 2007).

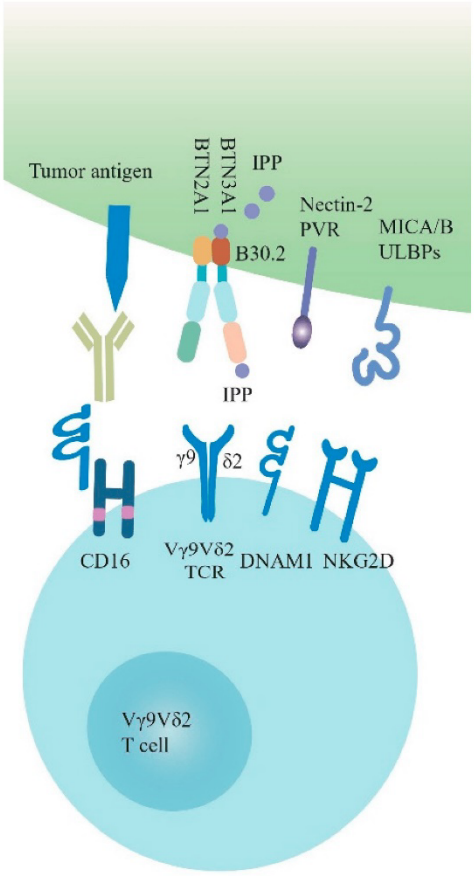


Figure 4.1.1: The identification of tumor cells by γδ T lymphocytes. γδ T cells are able to identify cancerous cells via γδ TCRs as well as NKRs. Vγ9Vδ2 TCR is dependent on BTN3A1 alongside BTN2A1 to identify the phosphoantigens (like IPP). Additionally, Vγ9Vδ2 T lymphocytes identify tumor cells by interacting with their respective ligands as shown. To activate Vγ9Vδ2-dependent ADCC, Vγ9Vδ2 T cells produce CD16, which recognizes different antibodies. Reproduced from Liu and Zhang, 2020.

A similar naturally occurring bacterial metabolite hydroxyl dimethylallyl pyrophosphate (HDMAPP, also known as HMBPP) is one of the strongest stimulants for Vγ9Vδ2 T cells (Vantourout et al., 2009). It has been shown that there is no absolute necessity of antigen presentation through APCs or antigen display via MHC for phosphoantigen mediated activation of Vγ9Vδ2 T cells (Bukowski et al., 1995; Morita et al., 2007). Butyrophillin (BTN) family members BTN3A1 and BTN2A1 play crucial roles in phosphoantigen sensing, activation, and proliferation of Vγ9Vδ2 cells (Karunakaran et al., 2020; Rigau et al., 2020).

The anti-tumor effect of γδ T cells is achieved by their virtue to produce pro-inflammatory cytokines interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α) which act in cohort with other factors to induce antitumor immunity as well as inhibition of cancer angiogenesis (Figure 4.1.2) (Li et al., 2008; Zhao et al., 2018). Activated γδ T cells also produce cytolytic proteins Granzyme B and Perforin, through which they lyse the tumor cells after migrating to the tumor microenvironment (Niu et al., 2015). In some tumors, upon hyperactivation of the mevalonate pathway, IPP is overproduced and activation of Vγ9Vδ2 T cells by IPP is dependent on the transmembrane butyrophillin molecules (Harly et al., 2012; Rigau et al., 2020; Sandstrom et al., 2014). Sensitivity of these tumors to lysis by Vγ9Vδ2 T cells increases upon treatment with aminobisphosphonates which leads to accumulation of intracellular IPP (Kunzmann et al., 2000; Kunzmann et al., 1999). We and others have previously shown that prior treatment of cancer cells with zoledronate, an aminobisphosphonate, can greatly increase the efficiency of lysis by activated Vγ9Vδ2 T cells (Dhar and Chiplunkar, 2010; Mattarollo et al., 2007).

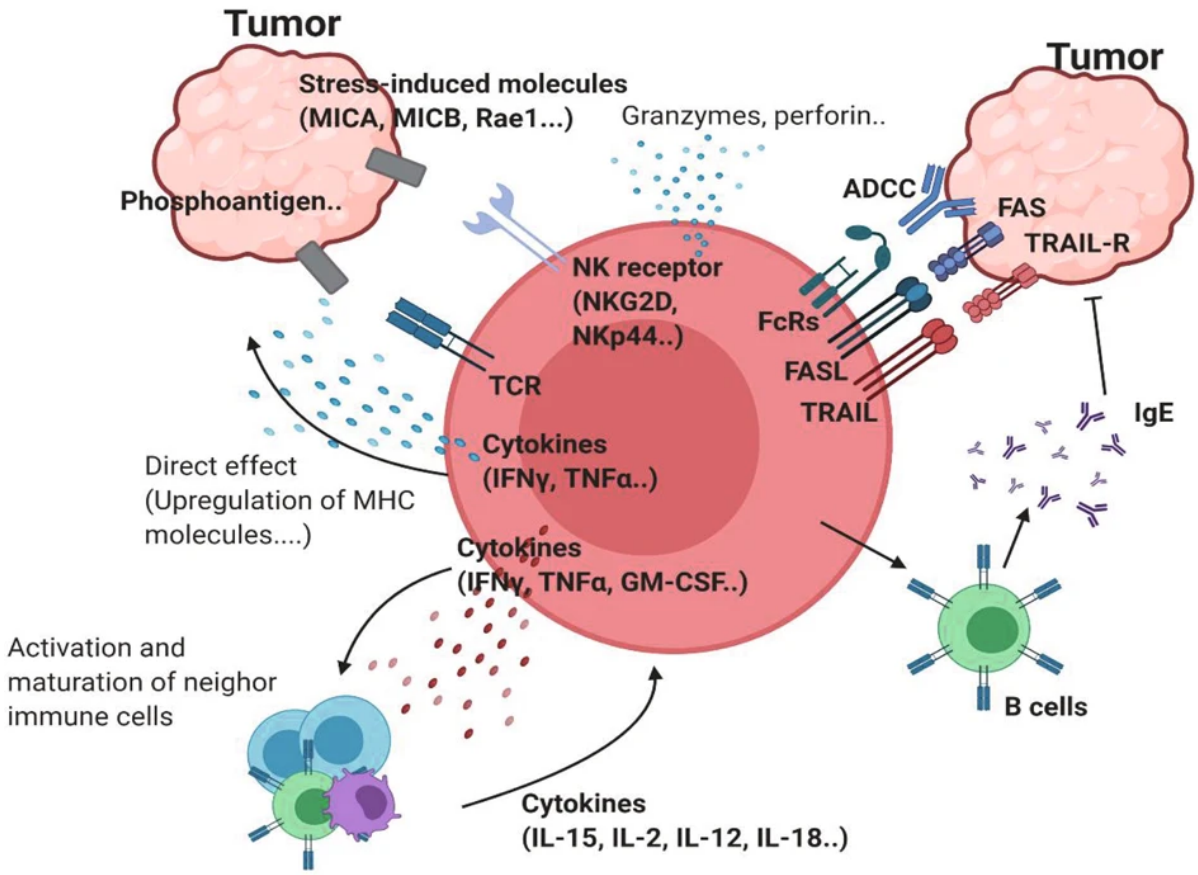


Figure 4.1.2: The anti-tumor effect of γδ T cells. Phosphoantigens (pAg) are γδTCR ligands that can attach to the receptor. MICA/B and Rae-1, two stress-triggered molecules, have the ability to attach onto the NK receptors like NKG2D. Protective γδ T cells get activated as a result of such ligation. Proinflammatory cytokines, including TNF or IFNγ, might influence other types of immune cells or induce MHC expression on the surface of tumor cells in order to enhance anti-tumor response. The tumor cells can be killed directly via the induction of cytotoxic proteins like Granzyme B and Perforin. Apart from these receptors, specific cytotoxicity targeting the tumor cells may additionally be induced by FcR-dependent, ADCC, and TRAIL recognition. IgE has anticancer properties and can be produced by B cells with the help of γδ T cells. Reproduced from Park and Lee, 2021.

Notch signaling has been extensively characterized in immune development and differentiation, as well as their maintenance and activation (Tanigaki et al., 2004; Yuan et al., 2010). It is essential for early T cell fate choice as well as αβ vs γδ lineage diversification. Notch signaling is also shown to promote anti-tumor activity of T cells and NK cells (Kelliher and Roderick, 2018). Our earlier results demonstrated that Notch expression in γδ T cells is

mediated by TCR activation, and inhibition of γ-secretase which cleaves Notch for nuclear export led to dramatic reduction in cytolytic activity of activated γδ T cells (Gogoi et al., 2014).

Multiple transcriptomics and genomics studies have provided insights of the spatio-temporal control of T cell activation, differentiation and development - especially in the context of $CD4^+$ and $CD8^+$ αβ subsets (Helgeland et al., 2020; Iwata et al., 2017; Rothenberg, 2014). Recently, numerous single-cell genomic studies have provided knowledge about the finer distinctions of T cell functionality at a single-cell resolution as well as the heterogeneity among marker-based sorted populations (Carmona et al., 2020; Efremova et al., 2020; Magen et al., 2019; Miller et al., 2019; Szabo et al., 2019). A comprehensive blood single-cell transcriptomics revealed that human TCR Vδ1 and TCR Vδ2 γδ T cells share cytotoxic hallmarks of both CD8 and NK cells, but form distinct clusters (Pizzolato et al., 2019).

Despite growing literature on the anti-tumor potential of γδ T cells (Silva-Santos et al., 2015), it is still unclear how activation via phosphoantigens or anti-CD3 antibodies mediates the effector functions of Vγ9Vδ2 T cells at the molecular level. Here, we performed RNA-sequencing (RNA-seq) of Vγ9Vδ2 T cells with multiple combinations of activating or repressive treatments and elucidated the primary transcriptional pathways utilized in each case. Our analyses revealed key transcription factors (TF) or their primary pathways that are affected via the activation/repression. This study provides important cues towards designing better combinations of target-specific molecules along with the current γδ T cell-based immunotherapies.

4.2 Materials and Methods

4.2.1 γδ T cell separation from peripheral blood

Blood samples were collected from three healthy volunteers and peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Hypaque (Sigma -Aldrich) differential density gradient centrifugation. The study was approved by the Institutional Ethics Committee and written informed consent was obtained from the healthy volunteers before collection of blood samples. Blood samples were processed as per the guidelines of the Institute Biosafety Committee. All experimental procedures involving clinical samples were handled in biosafety cabinets and laboratory personnel handling blood samples were vaccinated against Hepatitis B. γδ T cells were purified from PBMCs using anti-TCR γδ microbeads (clone 11F2; Miltenyi Biotech, Germany) by positive selection, as per the manufacturer's instructions.

4.2.2 Cell culture and treatments

Vγ9Vδ2 T cells were cultured in RPMI 1640 (Invitrogen) supplemented with 10% heat inactivated AB serum, 2 mM glutamine (Invitrogen), and Penicillin-Streptomycin (Sigma-Aldrich). Briefly, 1x106 Vγ9Vδ2 T cells, were seeded in triplicate sets in 24-well flat bottom plates (Nunc), and were treated in various combinations: unstimulated Vγ9Vδ2 T cells, Vγ9Vδ2 T cells stimulated with 50 IU/ml rIL2 (Peprotech), 50 IU/ml rIL2+HDMAPP (1 nM; Echelon), 50 IU/ml rIL2+IPP (40 μM; Sigma-Aldrich), 50 IU/ml rIL2+plate-bound anti-CD3 monoclonal antibody (clone OKT3; BD Biosciences, USA), rIL2+HDMAPP+γ-secretase inhibitor-X, L-685,458 (GSI-X, 15 μM) (Calbiochem, La Jolla, CA), rIL2+IPP+GSI-X, rIL2+anti-CD3+GSI-X, using previously standardized concentrations. After 72 h, the viability of Vγ9Vδ2 T cells was determined by Trypan Blue cell exclusion assay. The viability ranged from 86 to 90% for untreated Vγ9Vδ2 T cells and from 93.4 to 94.8% for all other treatments previously mentioned. The harvested cells were snap-frozen in TriZol reagent (Invitrogen) and stored at -80°C for library preparation.

4.2.3 RNA isolation, library preparation and sequencing

Total RNA was isolated using TriZol method. Quantitation of RNA was done on Qubit 4 Fluorometer (Thermo Scientific #Q33238) using RNA HS Kit (Thermo Scientific #Q32852). RIN value was checked using RNA IQ assay (Themo Scientific #Q33222) followed by running on 2100 Bioanalyzer (Agilent #G2939BA). All samples had a RIN value above 8. Five hundred

ng of total RNA was used to prepare libraries using TruSeq Stranded mRNA Sample Prep Kit (Illumina #20020594), with the TruSeq Stranded mRNA Sample Preparation Guide, Part #15031047. The libraries were sequenced using Illumina HiSeqX platform (Illumina, California, USA). The read length was 150 bp, paired end. Sequencing depth ranged from 40-70M reads.

4.2.4 Quality Control and read mapping

Initial quality of each sample was performed using FastQC v. 0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Low quality sequences and adapters were removed using Trimmomatic v. 0.39 (40). The following parameters were used for trimming: ILLUMINACLIP:TruSeq3-SE:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. More than 98% of paired-end reads survived after trimming in all samples. The reads were then mapped using the 'new tuxedo' package HISAT2 (41) to GRCh38 (hg38) genome. Annotation of properly oriented paired-reads to exons was done using FeatureCounts v. 2.0.1 (42) using these additional parameters: –s 2 –t exon –g gene id. The datasets presented in this study can be found in the GEO repository with accession number, GSE168642.

4.2.5 Differential gene expression (DGE) analysis

Differential gene expression analysis of count tables obtained from FeatureCounts (n=3, for each treatment) was performed using DESeq2 v.1.26.0 (43). To normalize the compositional variation in samples' libraries, we utilized the median of ratios method of the DESeq2 package. To obtain biologically relevant genes, for most analyses, we kept the DGE cutoff as logFC >0.5 or <1.5 and FDR <0.05. Enrichment of biological processes was performed using Gene Set Enrichment Analysis (http://software.broadinstitute.org/gsea) with GO terms obtained from MSigDB. For most significant pathway enrichment and comparative pathways across samples, we used ReactomePA v.1.30.0 package of R (44). Statistics was performed using Benjamini-Hochberg method with $p<0.01$.

Protein-Protein interaction maps were generated from the significant gene lists between treatments using Cytoscape v.3.8.0 (45) and Metascape (46). 2-dimensional cluster mapping of pathways was performed via ViSEAGO v.1.1.0 (47) package of R. Factor enrichment from DGE analysis was performed using EnrichR web server tool (48).

4.3 Results

To dissect the role of signaling pathways in the tumor microenvironment involved in the functioning/survival of Vγ9Vδ2 T cells, we employed ex-vivo culture of Vγ9Vδ2 T cells followed by high-throughput transcriptome analysis. The experimental regime is shown in Figure 4.3.1A.

4.3.1 IL2 mediated signaling reshapes Vγ9Vδ2 T cell gene-expression profile

Signaling via cytokine IL2 has been implicated in the survival as well as effector and memory functions of $CD4^+$ and $CD8^+$ αβ T cells (Ross and Cantrell, 2018). It has been shown that IL2 is important for γδ T cell survival and proliferation (Ribot et al., 2012). Additionally, γδ T cells produce IFN-γ in the presence of IL2. This occurs via de novo induction of T-bet and Eomesodermin (Eomes) by IL2 (Ribot et al., 2014). We, therefore, wished to dissect the global transcriptional changes that occur in Vγ9Vδ2 T cells upon IL2 treatment. We performed DGE analysis of IL2 treated vs untreated Vγ9Vδ2 T cells and mostly focused on only protein-coding transcripts with 38,557 genes annotated from the reference assembly. As shown via MA-plot, there were more than 1300 genes differentially regulated upon IL2 treatment with a stringent cutoff of >2.5 (logFC) (Figure 4.3.1B). We observed that about 71% of these genes were upregulated, suggesting that IL2 signal leads to better survival and effector function potentially via inducing expression of different signaling pathways. Furthermore, we performed hierarchical clustering for top 500 DE genes of both the datasets (Figure 4.3.1C). We observed that many genes amongst replicates do not show coherent expression. This could be due to varying state of IL2 signaling in the isolated pan γδ T cells from different donors. As shown, we found cell-cycle genes such as CDK1, CDK7 and AURKB as well as cytokine genes such as IL2RA, IL2, TNFSF8 and IL5 among others were upregulated, which is in line with the previously reported effect of IL2 on cultured Vγ9Vδ2 T cells (Vermijlen et al., 2007). We also performed gene-set enrichment analysis (GSEA) of DE genes between the two conditions using the Gene Ontology (GO)-term database. We observed that 91 GO terms were enriched using $p < 0.05$, and plotted both positively and negatively regulated processes with false discovery rate (FDR) cut-off as 0.05 (Figure 4.3.1D). Here, among the most significantly enriched processes were E2F mediated transcription, cell cycle, MTORC1 signaling and Wnt-β-catenin signaling axis. Further, multiple direct targets of Wnt/β-catenin signaling including DKK3, MMP9 and TCF7 were dysregulated. As expected, a large number of genes were differentially regulated in the presence of IL2, even with higher cut-offs for both p-values and logFC as depicted by the volcano plot (Figure 4.3.1E). From these observations, we determined the overall gene expression dynamics of Vγ9Vδ2 T cells upon IL2 signaling, and the pathways which get most affected.

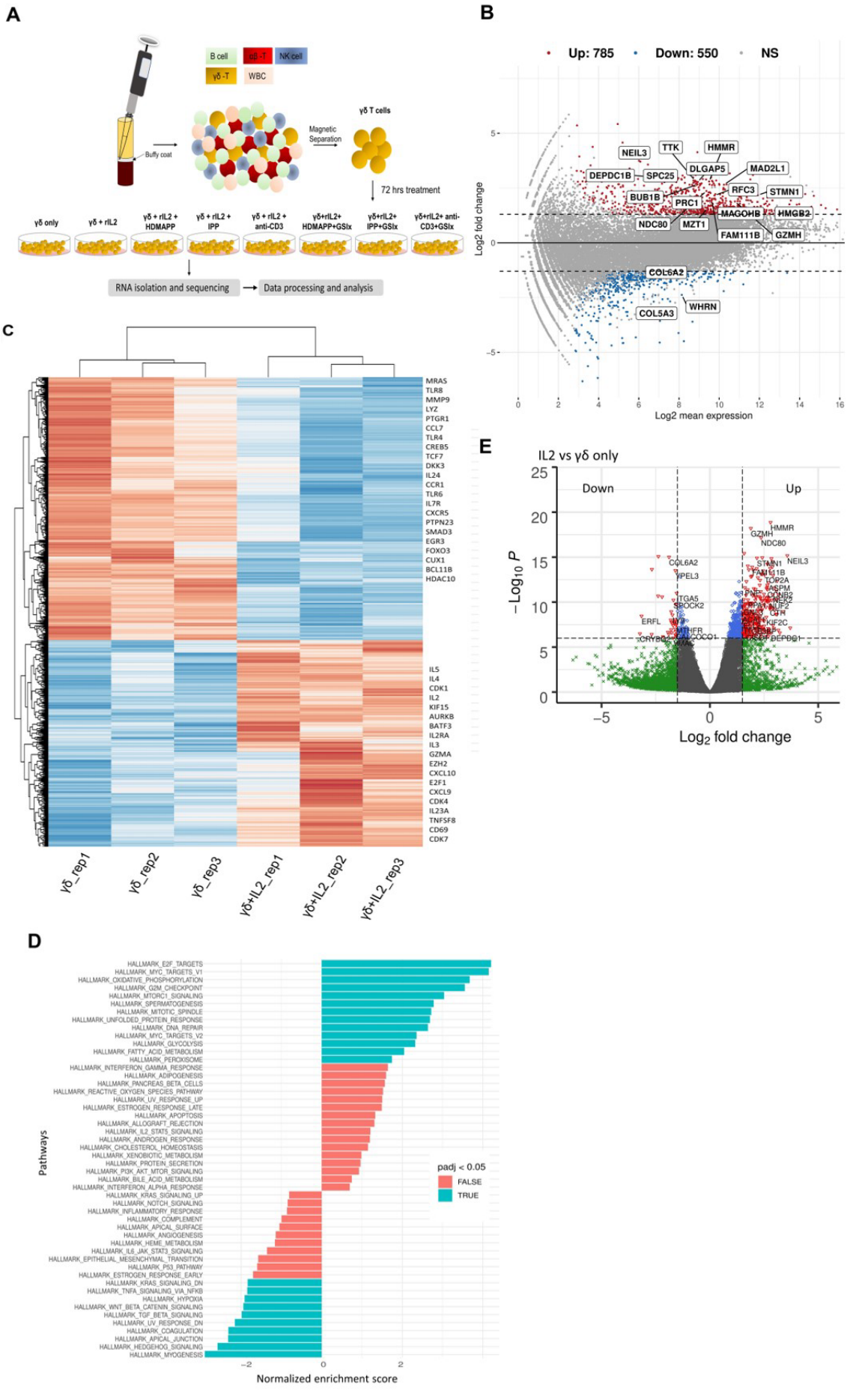

A
B cell
γδ -T
NK cell
γδ -T
WBC
Buffy coat
γδ T cells
Magnetic Separation
72 hrs treatment
γδ only
γδ + rIL2
γδ + rIL2 + HDMAPP
γδ + rIL2 + IPP
γδ + rIL2 + anti-CD3
γδ+rIL2+ HDMAPP+GSIx
γδ+rIL2+ IPP+GSIx
γδ+rIL2+ anti-CD3+GSIx
RNA isolation and sequencing
Data processing and analysis
B
Up: 785
Down: 550
NS
Log2 fold change
Log2 mean expression
NEIL3
TTK
HMMR
DEPDC1B
SPC25
DLGAP5
MAD2L1
RFC3
STMN1
BUB1B
PRC1
MAGOHB
HMGB2
NDC80
MZT1
FAM111B
GZMH
COL6A2
WHRN
COL5A3
C
MRAS
TLR8
MMP9
LY2
PTGR1
CCL7
TLR4
CREB5
TCF7
DKK3
IL24
CCR1
TLR6
IL7R
CXCR5
PTPN23
SMAD3
EGR3
FOXO3
CUX1
BCL11B
HDAC10
IL5
IL4
CDK1
IL2
KIF15
AURKB
BATF3
IL2RA
IL3
GZMA
EZH2
CXCL10
E2F1
CXCL9
CDK4
IL23A
TNFSF8
CD69
CDK7
γδ_rep1
γδ_rep2
γδ_rep3
γδ+IL2_rep1
γδ+IL2_rep2
γδ+IL2_rep3
D
HALLMARK_E2F_TARGETS
HALLMARK_MYC_TARGETS_V1
HALLMARK_OXIDATIVE_PHOSPHORYLATION
HALLMARK_G2M_CHECKPOINT
HALLMARK_MTORC1_SIGNALING
HALLMARK_SPERMATOGENESIS
HALLMARK_MITOTIC_SPINDLE
HALLMARK_UNFOLDED_PROTEIN_RESPONSE
HALLMARK_DNA_REPAIR
HALLMARK_MYC_TARGETS_V2
HALLMARK_GLYCOLYSIS
HALLMARK_FATTY_ACID_METABOLISM
HALLMARK_PEROXISOME
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_ADIPOGENESIS
HALLMARK_PANCREAS_BETA_CELLS
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY
HALLMARK_UV_RESPONSE_UP
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_APOPTOSIS
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_ANDROGEN_RESPONSE
HALLMARK_CHOLESTEROL_HOMEOSTASIS
HALLMARK_XENOBIOTIC_METABOLISM
HALLMARK_PROTEIN_SECRETION
HALLMARK_PI3K_AKT_MTOR_SIGNALING
HALLMARK_BILE_ACID_METABOLISM
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_NOTCH_SIGNALING
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_COMPLEMENT
HALLMARK_APICAL_SURFACE
HALLMARK_ANGIOGENESIS
HALLMARK_HEME_METABOLISM
HALLMARK_IL6_JAK_STAT3_SIGNALING
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_P53_PATHWAY
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_KRAS_SIGNALING_DN
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_HYPOXIA
HALLMARK_WNT_BETA_CATENIN_SIGNALING
HALLMARK_TGF_BETA_SIGNALING
HALLMARK_UV_RESPONSE_DN
HALLMARK_COAGULATION
HALLMARK_APICAL_JUNCTION
HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_MYOGENESIS
Pathways
padj < 0.05
FALSE
TRUE
Normalized enrichment score
E
IL2 vs γδ only
Down
Up
−Log10 P
Log2 fold change
HMMR
GZMH
NDC80
COL6A2
NEIL3
IPEL3
TGA5
SPOCK2
ERFL
KIF2C
DEPDC1

Figure 4.3.1: IL2 mediated signaling reshapes γδ T cell gene expression. A Schematic of γδ isolation, treatment regime and RNA-seq. **B** Overall gene-dysregulation upon IL2 treatment was plotted using MA function for fold-change vs normalized expression (n=3). Boxed genes shown are dysregulated by 2.5-logFC with an FDR of 0.05. **C** Clustered heatmap for top-most differentially regulated genes upon IL2 treatment. Key differentially regulated cytokine and effector genes are labelled. **D** Functional GeneSet Enrichment Analysis (fGSEA) was performed for Gene Ontology (GO) term or Hallmark enrichment terms. **E** Volcano plot for fold-change vs p-value. Intersecting lines indicate the logFC cutoff. Colors as indicated; Grey- NS, Blue- logP=-6, Green- FC=1.5, Red- FC=1.5 and logP=-6. Reproduced from Madhok et al., 2021.

4.3.2 Transcriptional dynamics upon phosphoantigen driven activation of Vγ9Vδ2 T cells

After determining how IL2 mediated responses of Vγ9Vδ2 T cells translate into transcriptional changes, we wanted to understand how non-peptide phoshoantigens IPP and HDMAPP affect the transcriptional program of Vγ9Vδ2 T cells. Our previous data demonstrated that treatment of γδ T cells with either IPP or a similar bacterial metabolite intermediate HDMAPP resulted in activation and increased cytotoxic potential in vitro (Bhat et al., 2018; Gogoi et al., 2014). We also showed that T-bet, Eomes and IFN-γ display a significant degree of upregulation post activation (Dhar and Chiplunkar, 2010; Gogoi et al., 2014). To understand the effect of phosphoantigen stimulation, we compared RNA-seq data of HDMAPP+IL2 treated Vγ9Vδ2 T cells to only IL2-treated cells as shown in Figure 4.3.2A. We observed an increase in the transcripts of pro-inflammatory cytokine receptor for TNF-α (TNFRS4) as well as AP1 family proteins FOS and JUN in HDMAPP+IL2 treated Vγ9Vδ2 T cells. Surprisingly, we observed only a mild but statistically insignificant (at $p<0.001$) increase in IFN-γ levels. Similarly, we analyzed IPP+IL2 vs IL2-only datasets and observed that expression of IL13 and CCR4 increased more than 2-fold (Figure 4.3.22B). We also observed an increase in the SLC family protein SLC7A5, an amino acid transporter involved in TCR activation of CD4 and CD8 T cells (Sinclair et al., 2013), in both the phosphoantigen treatments over IL2 only dataset. Although IL13 is known to be expressed exclusively by skin-residing γδ T cells displaying Vγ5Vδ1 TCR in mice model (Dalessandri et al., 2016), phosphoantigen treated human Vδ2 γδ T cells also secrete IL13 along with other cytokines (Peters et al., 2016). Since IPP and HDMAPP are similar metabolic intermediates in the presence of which γδ T cells seem to function similarly in vitro (Bhat et al., 2018), we expected to observe overlapping extent of differential gene

expression that these two treatments might confer when compared to IL2-only samples. As seen in Figure 2C, 61% and 50% overlap was observed for commonly upregulated genes in HDMAPP+IL2 vs IL2 and IPP+IL2 vs IL2 datasets respectively. In contrast, we observed 35% and 44% overlap between commonly downregulated genes among the two treatments. Given that many of the genes differentially regulated are not same in both up- and downregulated venn-sets in the two mentioned conditions, we wished to monitor which pathways are most significantly enriched for both HDMAPP and IPP treatments. In Figures 4.3.2D and 4.3.2E, we performed pathway analysis using enrich GO terms. We observed that most of the pathways with FDR<0.006 are different in both treatments, with lymphocyte specific pathways such as cytokine production and lymphocyte migration were enriched in the HDMAPP treatment. The cell-cycle and interferon-response pathways were most abundant in the IPP-treated Vγ9Vδ2 T cells. These results indicated that HDMAPP treatment can induce a robust immune response via activation of T-cell specific factors. On the other hand, IPP treatment seems to be more important in cell expansion as well as immune response.

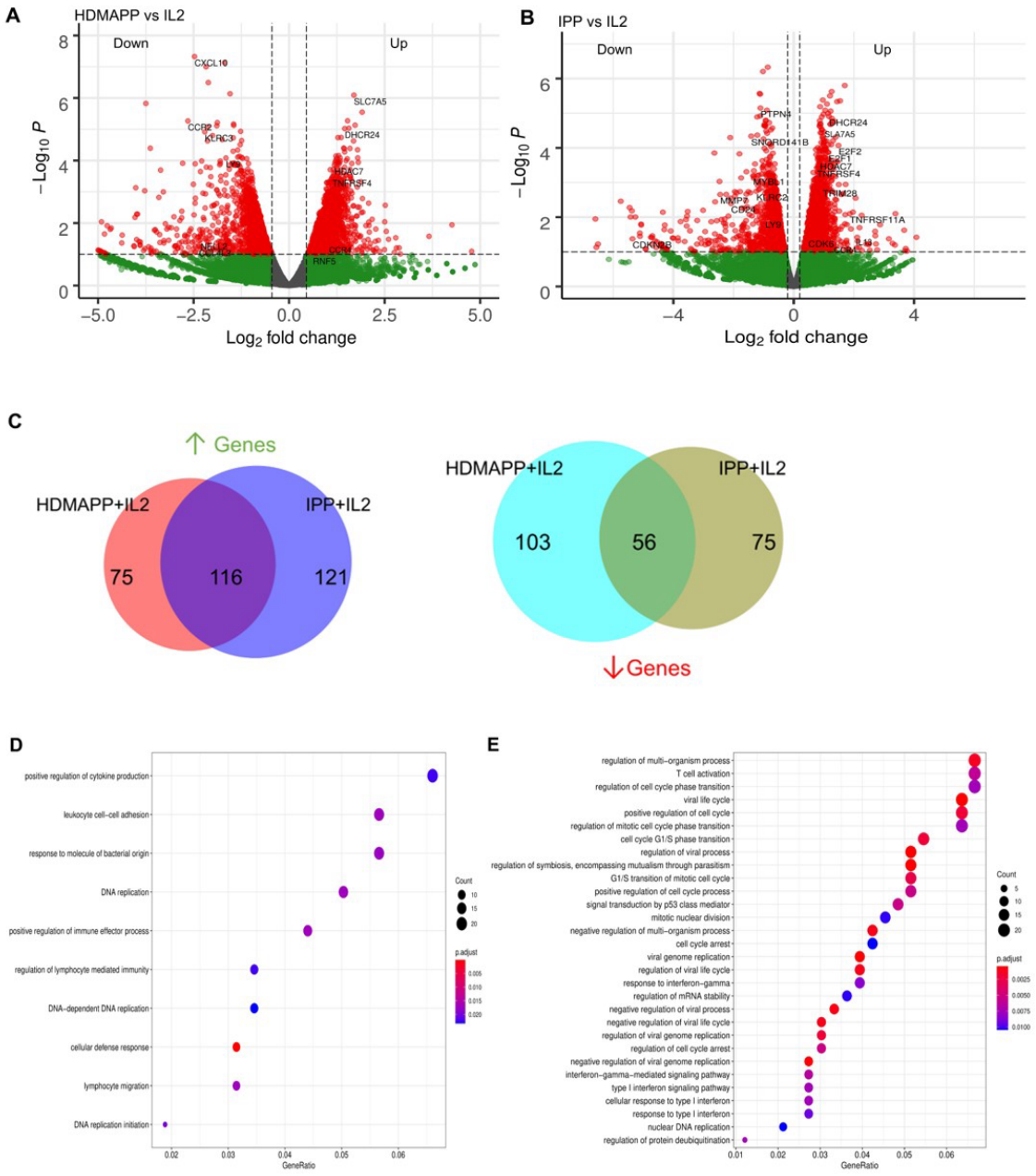


Figure 4.3.2: Transcriptional dynamics upon phosphoantigen driven activation of γδ T cells. A Differential gene expression of HDMAPP+IL2 vs IL2 treatment was performed showing key labels above a selected cutoff range. **B** Volcano plot for IPP+IL2 treatment compared for DGE over IL2. **C** Differentially regulated genes for HDMAPP+IL2 over IL2 and IPP+IL2 over IL2 with FDR<0.05 were segregated in up- and downregulated genes for each treatment pair. Venn-diagrams were plotted for common up- and down modulated genes comparing both the phosphoantigen conditions. **D** and **E** Enrichment dot plots for HDMAPP+IL2 over IL2 and IPP+IL2 over IL2 treatments, showing most significantly enriched GO terms. Reproduced from Madhok et al., 2021.

4.3.3 Distinct and common transcriptional pathways activated in Vγ9Vδ2 T cells upon stimulation with phosphoantigens

Based on the GO-term analysis (Figures 4.3.2D and 4.3.2E), we next studied the significantly enriched pathways in the two activation treatments. We performed Reactome pathway analysis to link the related pathways via network and plotted venn-diagram for common and unique pathways between HDMAPP and IPP stimulations (Figure 4.3.3A). We furthermore performed a Distance-mapping of Clustered GO terms in which we first clustered related GO-terms followed by computing their distance via semantic similarity with respect to each other on the XY axes. As shown in Figure 4.3.3B, most abundant clusters in HDMAPP treatment are of metabolic, stimuli responses and immune related processes. Of note, clusters of T cell activation and cell cycle have many overlapping GO terms, corroborating that these two 'umbrella' processes have mutual cross-talk in Vγ9Vδ2 T cells, which also supports the previously reported microarray-based gene expression profiling upon phosphoantigen BrHPP treatment (Pont et al., 2012). Interestingly, since we probed for the effects of only phosphoantigen treatment (HDMAPP and IPP) autonomous to IL2 influence, we observed pathways enriched only as a result of phosphoantigen treatments such as T cell activation and leukocyte differentiation. We furthermore probed into the TF and cytokine cross-talk via protein-protein interaction in the two treatments. Upon performing network analysis of interacting proteins with a threshold of >5 GO terms, many TFs and cytokines were enriched in HDMAPP (Figure 4.3.3C) and IPP (Figure 4.3.3D) DGEs over IL2 only datasets. Despite processes like 'chemokine signaling', 'leukocyte mediated cytotoxicity' and 'antigen presentation' being enriched in both conditions, they share only few TFs and chemokine receptors. It is also known that activation through anti-CD3 leads to γδ T cell activation and enhanced effector response (Gogoi et al., 2014), although there is a non-redundant requirement of anti-CD28 signal as well (Ribot et al., 2012). But since CD28 signaling activates IL2 mediated pathways (Ribot et al., 2012), we sought to compare dataset of anti-CD3+IL2 treated Vγ9Vδ2 T cells with other treatments. Hierarchical clustering of all four treatments- IL2 only, HDMAPP+IL2, IPP+IL2 and anti-CD3+IL2 was performed. We observed four major clusters based on DGE across conditions (Figure 4.3.3E). It is apparent that numerous top upregulated genes in anti-CD3 condition are not induced in IPP and HDMAPP-treated Vγ9Vδ2 T cells. We further performed the KEGG pathway comparison for the three activation treatments with FDR <0.05 (Figure 4.3.3F). Following the DGE profiles as shown in Figure 4.3.3E we observed that most significantly enriched pathways have only a few common elements such as 'cell cycle' and 'cytokine -chemokine signaling'. In terms of effector function, since majority of the γδ T cells cultured were of Vγ9Vδ2 type and produce IFN-γ upon phosphoantigen treatment (Bhat et al., 2018; Zhao et al., 2018), both IPP and HDMAPP

treatments showed enriched IFN signaling. Collectively, these results show that there seems to be both common and distinct metabolic as well as activation pathways that are induced upon phosphoantigen/anti-CD3 antibody treatment.

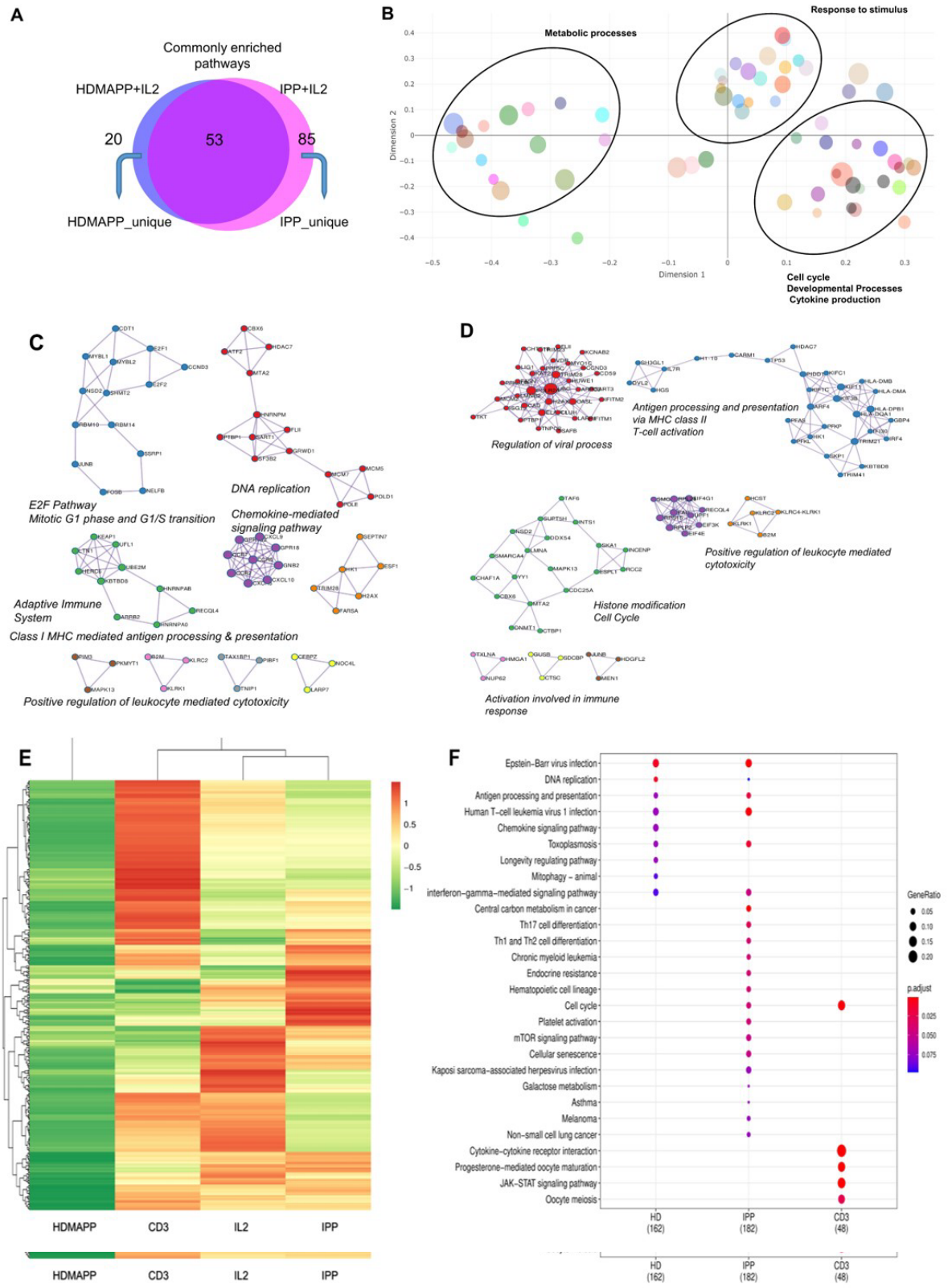


Figure 4.3.3: Distinct transcriptional pathways are activated via each phosphoantigen. **A** Intersection of significant GO terms enriched for HDMAPP and IPP treatments. **B** A

clustering-based semantic similarity mapping of most significant GO terms, showing how various pathways activated in HDMAPP treatment are distanced in a multi-dimensional scale. **C** and **D** Protein-protein interaction map for genes differentially enriched upon HDMAPP+IL2 and IPP+IL2 treatments over IL2, respectively. Most enriched corresponding processes are labelled. The maps were generated using Cytoscape with BioGrid, InWeb_IM and OmniPath databases. **E** Clustered heatmap of most differentially regulated genes with each mentioned stimulation. **F** KEGG pathway comparison was done for HDMAPP, IPP and anti-CD3 stimulations. The dots indicate both significant enrichment and p-value for the mentioned pathway. Reproduced from Madhok et al., 2021.

4.3.4 Inhibition of Notch signaling disrupts effector signaling of Vγ9Vδ2 T cells

Notch signaling plays an essential role in T cell activation and differentiation (Yuan et al., 2010), as well as their anti-tumor potential (Kelliher and Roderick, 2018). Our previously published data established that phosphoantigen or anti-CD3 driven in vitro activation of γδ T cells leads to induction of Notch signaling by directly regulating Notch expression (Gogoi et al., 2014), and blockade of Notch pathway via gamma secretase inhibitor (GSI-X) or siRNA results in decreased expression of effector molecules and hampered cytotoxic activity against tumor cell lines. To evaluate the global transcriptional phenomenon upon Notch inhibition in activated Vγ9Vδ2 T cells, we treated Vγ9Vδ2 T cells with a combination of either IPP+IL2+GSI, HDMAPP+IL2+GSI or anti-CD3+IL2+GSI and carried out RNA-seq. The datasets obtained were compared with their respective activated-only counterparts for DGE analysis. Notch inhibition resulted in greater number of downregulated genes compared to positive regulation for both phosphoantigens (cutoff >0.5logFC, FDR<0.05) (Figure 4.3.4A) as well as in the case of anti-CD3+IL2+GSI vs anti-CD3+IL2 treatments. As Notch signaling is known to feedback TCR signaling positively, we observed that inhibition of Notch signaling leads to downregulation of TCR-specific TFs such as BATF and RUNX2 as well as IL13 and CCR4 among cytokines and cytokine receptors (Figure 4.3.4B) while the top positively regulated genes included EGR3 and CD38 (Figure 4.3.4C). Although both BATF1 and 3 were downregulated commonly in all three inhibitory conditions, only BATF3 fulfilled our cutoff criterion. DGE analysis was also performed across activating and inhibitory conditions using the likelihood ratio test (similar to ANOVA) as shown via clustered heatmap (Figure 4.3.4D). Further, we also investigated which molecular pathways are most affected under Notch inhibition. We performed enrichment analysis using Enrichr (Chen et al., 2013) for IPP+GSI vs IPP DE genes (Figure 4.3.4E) and observed that the most affected pathways involve STAT6, T-BET and GATA3 mediated signaling. Strikingly, signaling via SATB1, an early TCR

responsive chromatin organizer (Patta et al., 2020), is also affected by Notch inhibition. In a similar analysis for HDMAPP+GSI vs HDMAPP dataset, we found that highest Enrichr scoring factors were STAT family proteins and NFATC2, both implicated in immune responses (Figure 4.3.4F). Interestingly, for anti-CD3+GSI we observed most dramatic changes in chromatin-related factors including, CEBP and EZH2 along with the STAT family proteins STAT6 and STAT3 (Figure 4.3.4G). Since cytokine signaling is important in the effector function of Vγ9Vδ2 T cells, we also plotted expression profiles of most DE cytokine and/or cytokine/chemokine receptor genes to represent the effect of Notch inhibition on activation of inflammatory response (Figure 4.3.4H). We observed an antagonistic effect of Notch inhibition on activation induced genes such as CCR4, IL5, IL13 and CXCL8. These results collectively point out that there are distinct TF networks enriched upon different activation treatments thus modulating the effector function of Vγ9Vδ2 T cells uniquely. Our results also indicate that inhibition of Notch signaling disrupts the phosphoantigen-mediated effector gene-signature of Vγ9Vδ2 T cells hampering their function.

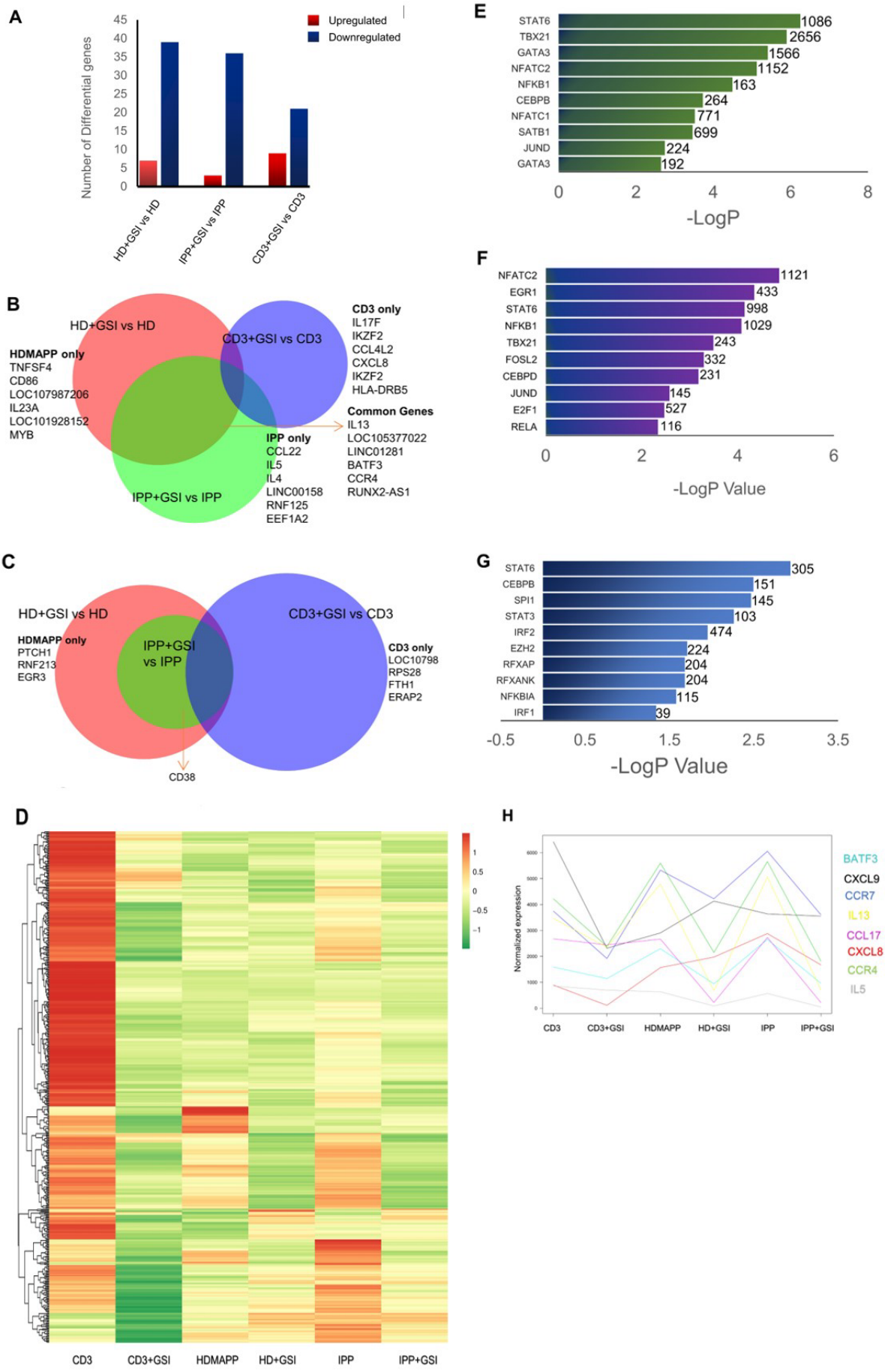

A
Upregulated
Downregulated
Number of Differential genes
HD+GSI vs HD
IPP+GSI vs IPP
CD3+GSI vs CD3
B
HD+GSI vs HD
CD3+GSI vs CD3
IPP+GSI vs IPP
HDMAPP only
TNFSF4
CD86
LOC107987206
IL23A
LOC101928152
MYB
CD3 only
IL17F
IKZF2
CCL4L2
CXCL8
IKZF2
HLA-DRB5
Common Genes
IL13
LOC105377022
LINC01281
BATF3
CCR4
RUNX2-AS1
IPP only
CCL22
IL5
IL4
LINC00158
RNF125
EEF1A2
C
HD+GSI vs HD
CD3+GSI vs CD3
IPP+GSI vs IPP
HDMAPP only
PTCH1
RNF213
EGR3
CD3 only
LOC10798
RPS28
FTH1
ERAP2
CD38
D
CD3
CD3+GSI
HDMAPP
HD+GSI
IPP
IPP+GSI
E
STAT6 1086
TBX21 2656
GATA3 1566
NFATC2 1152
NFKB1 163
CEBPB 264
NFATC1 771
SATB1 699
JUND 224
GATA3 192
-LogP
F
NFATC2 1121
EGR1 433
STAT6 998
NFKB1 1029
TBX21 243
FOSL2 332
CEBPD 231
JUND 145
E2F1 527
RELA 116
-LogP Value
G
STAT6 305
CEBPB 151
SPI1 145
STAT3 103
IRF2 474
EZH2 224
RFXAP 204
RFXANK 204
NFKBIA 115
IRF1 39
-LogP Value
H
Normalized expression
BATF3
CXCL9
CCR7
IL13
CCL17
CXCL8
CCR4
IL5
CD3
CD3+GSI
HDMAPP
HD+GSI
IPP
IPP+GSI

Figure 4.3.4: Inhibition of Notch Signaling disrupts effector signaling of γδ T cells. A Total number of up- and downregulated genes in each activation vs corresponding activation + Notch-inhibition condition. DGE was performed with Wald's test, logFC>0.5 and FDR<0.05. **B** and **C** Common genes down- and upregulated between each Notch inhibited dataset. Some of the key factors dysregulated are shown for both overlapping and individual treatments. **D** Heatmap showing differential clustering of γδ T cell activating treatments and their combination with Notch inhibition. **E**, **F** and **G** Enrichr analysis was performed on each DGE analyzed pair – IPP+GSI vs IPP, HDMAPP+GSI vs HD and anti-CD3+GSI vs CD3, respectively. The most significantly affected TFs mediated pathways are shown, ordered according to their -logP values. Combined Enrichr score is indicated for each factor as shown. **H** Normalized gene expression profiles for key cytokines across different activations and activations coupled with Notch inhibition. Statistics was performed using the Likelihood ratio test, with FDR<0.05. Reproduced from Madhok et al., 2021.

4.4 Discussion

Present study reveals the transcriptional dynamics during γδ T cell activation. We showed that stimulations of TCR using phosphoantigen (IPP and HDMAPP) and anti-CD3 antibody treatments in vitro can lead to induction of both distinct and similar transcriptional programs in Vγ9Vδ2 T cells, although the effector molecules such as IFN-γ, PRF1 and TNFRSF members are positively regulated across various activation methods. It has been reported that HDMAPP is 1000-fold more potent than IPP in activating Vγ9Vδ2 T cells. Binding studies using isothermal titration calorimetry have demonstrated that HDMAPP binds to B30.2 intracellular domain of BTN3A1 with an affinity of 0.5 μM whereas IPP bound with 0.5 mM affinity reflecting the potency differences between these two agonists in activating γδ T cells (Sandstrom et al., 2014). HMBPP stimulation of γδ T cells induces formation of high-density nanoclusters of Vγ9Vδ2 TCR on the membranes of these cells, and maintain high TCR surface expression in contrast to extensive internalization of the TCR observed after anti-CD3 stimulation (Chen et al., 2008; Correia et al., 2009; Sureshbabu et al., 2020). There are fundamental differences in activation of γδ T cells by anti-CD3 and phosphoantigens. Phosphoantigens do not trigger the CD3 conformational change in γδ TCR (Dopfer et al., 2014). Anti-CD3 activation enhances the recruitment of non-catalytic region of tyrosine kinase adapter protein 1 (NCK) to the γδ TCR which helps in stabilizing the receptor in an active conformation (Juraske et al., 2018). Interestingly, Notch signaling has been reported to be involved in the activation of Th17 γδ T cells as well. IL17 and RORγT are direct transcriptional targets of Notch signaling (Keerthivasan et al., 2011) and Notch-Hes1 pathway has been shown to be involved in the murine intrathymic development of Th17 γδ T cells (Papotto et al., 2017; Shibata et al., 2011). Similar to our observation, overexpression of CCR4 receptor has been reported on γδ T cells isolated from peripheral blood of healthy individuals upon stimulation with IPP (Brandes et al., 2003). Elevated expression of CCR4 receptor on human Vδ1 (Zhou et al., 2012) and Vδ2 subsets (Yin et al., 2015) has also been reported under certain pathological conditions. The increase in the CCR4 levels on activated γδ T cells has been attributed to the migratory ability of these cells.

Recently, single-cell transcriptomic analysis revealed the hallmarks of distinction and overlap between different γδ T cells isolated from peripheral blood (Efremova et al., 2020). Antigen-stimulation in the periphery leads to specific expansion of Vγ9Vδ2 T cells throughout the post-natal life (Ravens et al., 2017). Thus, this subset of γδ T cells has been at the center-stage of γδ T cell-based therapies (Fowler and Bodman-Smith, 2015). We showed that distinct

transcription factors/pathways are activated contingent on the kind of antigen. To our surprise, IFN-signaling was commonly enriched upon IPP and HDMAPP stimulation but not with anti-CD3, although IFN-γ levels increased modestly over IL2 treatment. Earlier gene expression data on TCR associated transcriptional signatures of cytokines supports high IFN-γ and low IL17 expression in γδ T cells stimulated with HMBPP. In previous studies, microarray data for differential gene expression of HMB-PP+IL2 and 'resting' γδ T cells identified many key genes involved in proliferation and cytotoxic activity of Vγ9Vδ2 T cells (Correia et al., 2009; Pont et al., 2012). Many of these genes are enriched in our datasets as well like MAPK13, SDF4, MAZ and LAG3. Silva-Santos and colleagues have demonstrated the importance of PI-3K/AKT and ERK/MAPK signaling for Vγ9Vδ2 proliferation and cytokine release (Correia et al., 2009). Further experiments are warranted to see the effect of factor-specific ablation such as BATF3 and NFkB on Vγ9Vδ2 T cell activation.

Previous studies by Chiplunkar's group have demonstrated an obligatory role of Notch signaling on Vγ9Vδ2 T cell effector function (Bhat et al., 2018; Gogoi et al., 2014). It has been shown using microarray data that inhibition of Notch signaling resulted in metabolic stress as well as activation deficit in B cell precursors (Meng et al., 2011), although no such global studies in Vγ9Vδ2 subsets are done to our knowledge. To predict the mechanism of Notch mediated transcriptional dynamics in TCR sensitized Vγ9Vδ2 T cells, we utilized a similar inhibitory approach combined with antigenic or antibody treatments. Our RNA-seq analysis not only corroborated that effector molecules such as IFN-γ and Granzyme B are downregulated upon Notch blockade, but also provided a large set of potential targets to explore in the future studies. Many clinical studies have attempted to harness the anti-tumor activity of γδ T cells through isolation of Vγ9Vδ2 T cells, their ex vivo expansion and phosphoantigen activation, and adoptive transfer in patients (Bennouna et al., 2010; Meraviglia et al., 2010; Wilhelm et al., 2003). Although there were no side-effects, the success in these clinical studies was not substantial. This is attributed to the deficits in the knowledge of Vγ9Vδ2 TCR interaction dynamics with the stimulating antigen as well as the functional outset of such interaction (Sebestyen et al., 2020). Our global transcriptome analysis pertaining to such in vitro activation treatments displays the range of pathways and highlights the key molecules affected therein. The functional role of these pathways/TFs needs to be ascertained with respect to anti-tumor cytotoxicity of Vγ9Vδ2 T cells, and requires future experimentation. Upon phosphoantigen activation combined with Notch inhibition, one of the factors enriched in our regulatory analysis was SATB1, which suggests that genes directly regulated by SATB1 are responsive to both TCR and Notch signaling.

In conclusion, we demonstrated that effector functions of Vγ9Vδ2 T cells from peripheral blood ensue from many immunologic and metabolic pathways, and these pathways are disrupted in consequence to Notch signaling inhibition. The specific role of treatment-dependent factors can be altered via CRISPR/Cas9 methodology followed by antigen stimulation, which might lead to an improvement in anti-cancer potential of peripheral Vγ9Vδ2 T cells. To link epigenetic changes to transcriptional dynamics upon these treatments, we have also performed ChIP-seq of histone marks such as H3K4me3, H3K27ac and H3K9me3 in a similarly treated Vγ9Vδ2 T cells. These data will be integrated with the transcriptomics analysis for more comprehensive understanding of Vγ9Vδ2 T cell's functionality as part of future studies.

4.5 References

Bennouna, J., Levy, V., Sicard, H., Senellart, H., Audrain, M., Hiret, S., Rolland, F., Bruzzoni-Giovanelli, H., Rimbert, M., Galea, C*., et al.* (2010). Phase I study of bromohydrin pyrophosphate (BrHPP, IPH 1101), a Vgamma9Vdelta2 T lymphocyte agonist in patients with solid tumors. Cancer Immunol Immunother *59*, 1521-1530.

Bhat, S.A., Vedpathak, D.M., and Chiplunkar, S.V. (2018). Checkpoint Blockade Rescues the Repressive Effect of Histone Deacetylases Inhibitors on gammadelta T Cell Function. Front Immunol *9*, 1615.

Brandes, M., Willimann, K., Lang, A.B., Nam, K.H., Jin, C., Brenner, M.B., Morita, C.T., and Moser, B. (2003). Flexible migration program regulates gamma delta T-cell involvement in humoral immunity. Blood *102*, 3693-3701.

Bukowski, J.F., Morita, C.T., Tanaka, Y., Bloom, B.R., Brenner, M.B., and Band, H. (1995). V gamma 2V delta 2 TCR-dependent recognition of non-peptide antigens and Daudi cells analyzed by TCR gene transfer. J Immunol *154*, 998-1006.

Carmona, S.J., Siddiqui, I., Bilous, M., Held, W., and Gfeller, D. (2020). Deciphering the transcriptomic landscape of tumor-infiltrating CD8 lymphocytes in B16 melanoma tumors with single-cell RNA-Seq. Oncoimmunology *9*, 1737369.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Chen, Y., Shao, L., Ali, Z., Cai, J., and Chen, Z.W. (2008). NSOM/QD-based nanoscale immunofluorescence imaging of antigen-specific T-cell receptor responses during an in vivo clonal Vgamma2Vdelta2 T-cell expansion. Blood *111*, 4220-4232.

Chiplunkar, S., Dhar, S., Wesch, D., and Kabelitz, D. (2009). gammadelta T cells in cancer immunotherapy: current status and future prospects. Immunotherapy *1*, 663-678.

Constant, P., Davodeau, F., Peyrat, M.A., Poquet, Y., Puzo, G., Bonneville, M., and Fournie, J.J. (1994). Stimulation of human gamma delta T cells by nonpeptidic mycobacterial ligands. Science *264*, 267-270.

Correia, D.V., d'Orey, F., Cardoso, B.A., Lanca, T., Grosso, A.R., deBarros, A., Martins, L.R., Barata, J.T., and Silva-Santos, B. (2009). Highly active microbial phosphoantigen induces rapid yet sustained MEK/Erk- and PI-3K/Akt-mediated signal transduction in anti-tumor human gammadelta T-cells. PLoS One *4*, e5657.

Dalessandri, T., Crawford, G., Hayes, M., Castro Seoane, R., and Strid, J. (2016). IL-13 from intraepithelial lymphocytes regulates tissue homeostasis and protects against carcinogenesis in the skin. Nat Commun *7*, 12080.

Dar, A.A., Patil, R.S., and Chiplunkar, S.V. (2014). Insights into the Relationship between Toll Like Receptors and Gamma Delta T Cell Responses. Front Immunol *5*, 366.

Dhar, S., and Chiplunkar, S.V. (2010). Lysis of aminobisphosphonate-sensitized MCF-7 breast tumor cells by Vgamma9Vdelta2 T cells. Cancer Immun *10*, 10.

Dopfer, E.P., Hartl, F.A., Oberg, H.H., Siegers, G.M., Yousefi, O.S., Kock, S., Fiala, G.J., Garcillan, B., Sandstrom, A., Alarcon, B.*, et al.* (2014). The CD3 conformational change in the gammadelta T cell receptor is not triggered by antigens but can be enforced to enhance tumor killing. Cell Rep *7*, 1704-1715.

Efremova, M., Vento-Tormo, R., Park, J.E., Teichmann, S.A., and James, K.R. (2020). Immunology in the Era of Single-Cell Technologies. Annu Rev Immunol *38*, 727-757.

Fowler, D.W., and Bodman-Smith, M.D. (2015). Harnessing the power of Vdelta2 cells in cancer immunotherapy. Clin Exp Immunol *180*, 1-10.

Gogoi, D., Dar, A.A., and Chiplunkar, S.V. (2014). Involvement of Notch in activation and effector functions of gammadelta T cells. J Immunol *192*, 2054-2062.

Harly, C., Guillaume, Y., Nedellec, S., Peigne, C.M., Monkkonen, H., Monkkonen, J., Li, J., Kuball, J., Adams, E.J., Netzer, S.*, et al.* (2012). Key implication of CD277/butyrophilin-3 (BTN3A) in cellular stress sensing by a major human gammadelta T-cell subset. Blood *120*, 2269-2279.

Hayday, A.C. (2000). [gamma][delta] cells: a right time and a right place for a conserved third way of protection. Annu Rev Immunol *18*, 975-1026.

Hayday, A.C., Saito, H., Gillies, S.D., Kranz, D.M., Tanigawa, G., Eisen, H.N., and Tonegawa, S. (1985). Structure, organization, and somatic rearrangement of T cell gamma genes. Cell *40*, 259-269.

Helgeland, H., Gabrielsen, I., Akselsen, H., Sundaram, A.Y.M., Flam, S.T., and Lie, B.A. (2020). Transcriptome profiling of human thymic CD4+ and CD8+ T cells compared to primary peripheral T cells. BMC Genomics *21*, 350.

Holtmeier, W., and Kabelitz, D. (2005). gammadelta T cells link innate and adaptive immune responses. Chem Immunol Allergy *86*, 151-183.

Iwata, A., Durai, V., Tussiwand, R., Briseno, C.G., Wu, X., Grajales-Reyes, G.E., Egawa, T., Murphy, T.L., and Murphy, K.M. (2017). Quality of TCR signaling determined by differential affinities of enhancers for the composite BATF-IRF4 transcription factor complex. Nat Immunol *18*, 563-572.

Juraske, C., Wipa, P., Morath, A., Hidalgo, J.V., Hartl, F.A., Raute, K., Oberg, H.H., Wesch, D., Fisch, P., Minguet, S.*, et al.* (2018). Anti-CD3 Fab Fragments Enhance Tumor Killing by Human gammadelta T Cells Independent of Nck Recruitment to the gammadelta T Cell Antigen Receptor. Front Immunol *9*, 1579.

Kadivar, M., Petersson, J., Svensson, L., and Marsal, J. (2016). CD8alphabeta+ gammadelta T Cells: A Novel T Cell Subset with a Potential Role in Inflammatory Bowel Disease. J Immunol *197*, 4584-4592.

Karunakaran, M.M., Willcox, C.R., Salim, M., Paletta, D., Fichtner, A.S., Noll, A., Starick, L., Nohren, A., Begley, C.R., Berwick, K.A.*, et al.* (2020). Butyrophilin-2A1 Directly Binds Germline-Encoded Regions of the Vgamma9Vdelta2 TCR and Is Essential for Phosphoantigen Sensing. Immunity *52*, 487-498 e486.

Keerthivasan, S., Suleiman, R., Lawlor, R., Roderick, J., Bates, T., Minter, L., Anguita, J., Juncadella, I., Nickoloff, B.J., Le Poole, I.C.*, et al.* (2011). Notch signaling regulates mouse and human Th17 differentiation. J Immunol *187*, 692-701.

Kelliher, M.A., and Roderick, J.E. (2018). NOTCH Signaling in T-Cell-Mediated Anti-Tumor Immunity and T-Cell-Based Immunotherapies. Front Immunol *9*, 1718.

Kunzmann, V., Bauer, E., Feurle, J., Weissinger, F., Tony, H.P., and Wilhelm, M. (2000). Stimulation of gammadelta T cells by aminobisphosphonates and induction of antiplasma cell activity in multiple myeloma. Blood *96*, 384-392.

Kunzmann, V., Bauer, E., and Wilhelm, M. (1999). Gamma/delta T-cell stimulation by pamidronate. N Engl J Med *340*, 737-738.

Lang, F., Peyrat, M.A., Constant, P., Davodeau, F., David-Ameline, J., Poquet, Y., Vie, H., Fournie, J.J., and Bonneville, M. (1995). Early activation of human V gamma 9V delta 2 T cell broad cytotoxicity and TNF production by nonpeptidic mycobacterial ligands. J Immunol *154*, 5986-5994.

Li, H., Luo, K., and Pauza, C.D. (2008). TNF-alpha is a positive regulatory factor for human Vgamma2 Vdelta2 T cells. J Immunol *181*, 7131-7137.

Liu, Y., and Zhang, C. (2020). The Role of Human gammadelta T Cells in Anti-Tumor Immunity and Their Potential for Cancer Immunotherapy. Cells *9(5):1206*.

Madhok, A., Bhat, S.A., Philip, C.S., Sureshbabu, S.K., Chiplunkar, S., and Galande, S. (2021). Transcriptome Signature of Vgamma9Vdelta2 T Cells Treated With Phosphoantigens and Notch Inhibitor Reveals Interplay Between TCR and Notch Signaling Pathways. Front Immunol *12*, 660361.

Magen, A., Nie, J., Ciucci, T., Tamoutounour, S., Zhao, Y., Mehta, M., Tran, B., McGavern, D.B., Hannenhalli, S., and Bosselut, R. (2019). Single-Cell Profiling Defines Transcriptomic Signatures Specific to Tumor-Reactive versus Virus-Responsive CD4(+) T Cells. Cell Rep *29*, 3019-3032 e3016.

Mattarollo, S.R., Kenna, T., Nieda, M., and Nicol, A.J. (2007). Chemotherapy and zoledronate sensitize solid tumour cells to Vgamma9Vdelta2 T cell cytotoxicity. Cancer Immunol Immunother *56*, 1285-1297.

Meng, X., Matlawska-Wasowska, K., Girodon, F., Mazel, T., Willman, C.L., Atlas, S., Chen, I.M., Harvey, R.C., Hunger, S.P., Ness, S.A.*, et al.* (2011). GSI-I (Z-LLNle-CHO) inhibits gamma-secretase and the proteosome to trigger cell death in precursor-B acute lymphoblastic leukemia. Leukemia *25*, 1135-1146.

Meraviglia, S., Eberl, M., Vermijlen, D., Todaro, M., Buccheri, S., Cicero, G., La Mendola, C., Guggino, G., D'Asaro, M., Orlando, V.*, et al.* (2010). In vivo manipulation of Vgamma9Vdelta2 T cells with zoledronate and low-dose interleukin-2 for immunotherapy of advanced breast cancer patients. Clin Exp Immunol *161*, 290-297.

Miller, B.C., Sen, D.R., Al Abosy, R., Bi, K., Virkud, Y.V., LaFleur, M.W., Yates, K.B., Lako, A., Felt, K., Naik, G.S.*, et al.* (2019). Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade. Nat Immunol *20*, 326-336.

Morita, C.T., Jin, C., Sarikonda, G., and Wang, H. (2007). Nonpeptide antigens, presentation mechanisms, and immunological memory of human Vgamma2Vdelta2 T cells: discriminating friend from foe through the recognition of prenyl pyrophosphate antigens. Immunol Rev *215*, 59-76.

Niu, C., Jin, H., Li, M., Xu, J., Xu, D., Hu, J., He, H., Li, W., and Cui, J. (2015). In vitro analysis of the proliferative capacity and cytotoxic effects of ex vivo induced natural killer cells, cytokine-induced killer cells, and gamma-delta T cells. BMC Immunol *16*, 61.

Papotto, P.H., Goncalves-Sousa, N., Schmolka, N., Iseppon, A., Mensurado, S., Stockinger, B., Ribot, J.C., and Silva-Santos, B. (2017). IL-23 drives differentiation of peripheral

gammadelta17 T cells from adult bone marrow-derived precursors. EMBO Rep *18*, 1957-1967.

Park, J.H., and Lee, H.K. (2021). Function of gammadelta T cells in tumor immunology and their application to cancer therapy. Exp Mol Med *53*, 318-327.

Patta, I., Madhok, A., Khare, S., Gottimukkala, K.P., Verma, A., Giri, S., Dandewad, V., Seshadri, V., Lal, G., Misra-Sen, J.*, et al.* (2020). Dynamic regulation of chromatin organizer SATB1 via TCR-induced alternative promoter switch during T-cell development. Nucleic Acids Res *48*, 5873-5890.

Peters, C., Hasler, R., Wesch, D., and Kabelitz, D. (2016). Human Vdelta2 T cells are a major source of interleukin-9. Proc Natl Acad Sci U S A *113*, 12520-12525.

Pizzolato, G., Kaminski, H., Tosolini, M., Franchini, D.M., Pont, F., Martins, F., Valle, C., Labourdette, D., Cadot, S., Quillet-Mary, A.*, et al.* (2019). Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRVdelta1 and TCRVdelta2 gammadelta T lymphocytes. Proc Natl Acad Sci U S A *116*, 11906-11915.

Pont, F., Familiades, J., Dejean, S., Fruchon, S., Cendron, D., Poupot, M., Poupot, R., L'Faqihi-Olive, F., Prade, N., Ycart, B.*, et al.* (2012). The gene expression profile of phosphoantigen-specific human gammadelta T lymphocytes is a blend of alphabeta T-cell and NK-cell signatures. Eur J Immunol *42*, 228-240.

Ravens, S., Schultze-Florey, C., Raha, S., Sandrock, I., Drenker, M., Oberdorfer, L., Reinhardt, A., Ravens, I., Beck, M., Geffers, R.*, et al.* (2017). Human gammadelta T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. Nat Immunol *18*, 393-401.

Ribot, J.C., Debarros, A., Mancio-Silva, L., Pamplona, A., and Silva-Santos, B. (2012). B7-CD28 costimulatory signals control the survival and proliferation of murine and human gammadelta T cells via IL-2 production. J Immunol *189*, 1202-1208.

Ribot, J.C., Ribeiro, S.T., Correia, D.V., Sousa, A.E., and Silva-Santos, B. (2014). Human gammadelta thymocytes are functionally immature and differentiate into cytotoxic type 1 effector T cells upon IL-2/IL-15 signaling. J Immunol *192*, 2237-2243.

Rigau, M., Ostrouska, S., Fulford, T.S., Johnson, D.N., Woods, K., Ruan, Z., McWilliam, H.E.G., Hudson, C., Tutuka, C., Wheatley, A.K.*, et al.* (2020). Butyrophilin 2A1 is essential for phosphoantigen reactivity by gammadelta T cells. Science *367, eaay5516*.

Ross, S.H., and Cantrell, D.A. (2018). Signaling and Function of Interleukin-2 in T Lymphocytes. Annu Rev Immunol *36*, 411-433.

Rothenberg, E.V. (2014). The chromatin landscape and transcription factors in T cell programming. Trends Immunol *35*, 195-204.

Sandstrom, A., Peigne, C.M., Leger, A., Crooks, J.E., Konczak, F., Gesnel, M.C., Breathnach, R., Bonneville, M., Scotet, E., and Adams, E.J. (2014). The intracellular B30.2 domain of butyrophilin 3A1 binds phosphoantigens to mediate activation of human Vgamma9Vdelta2 T cells. Immunity *40*, 490-500.

Sebestyen, Z., Prinz, I., Dechanet-Merville, J., Silva-Santos, B., and Kuball, J. (2020). Translating gammadelta (gammadelta) T cells and their receptors into cancer cell therapies. Nat Rev Drug Discov *19*, 169-184.

Shibata, K., Yamada, H., Sato, T., Dejima, T., Nakamura, M., Ikawa, T., Hara, H., Yamasaki, S., Kageyama, R., Iwakura, Y*., et al.* (2011). Notch-Hes1 pathway is required for the development of IL-17-producing gammadelta T cells. Blood *118*, 586-593.

Silva-Santos, B., Serre, K., and Norell, H. (2015). gammadelta T cells in cancer. Nat Rev Immunol *15*, 683-691.

Sinclair, L.V., Rolf, J., Emslie, E., Shi, Y.B., Taylor, P.M., and Cantrell, D.A. (2013). Control of amino-acid transport by antigen receptors coordinates the metabolic reprogramming essential for T cell differentiation. Nat Immunol *14*, 500-508.

Sureshbabu, S.K., Chaukar, D., and Chiplunkar, S.V. (2020). Hypoxia regulates the differentiation and anti-tumor effector functions of gammadeltaT cells in oral cancer. Clin Exp Immunol *201*, 40-57.

Szabo, P.A., Levitin, H.M., Miron, M., Snyder, M.E., Senda, T., Yuan, J., Cheng, Y.L., Bush, E.C., Dogra, P., Thapa, P*., et al.* (2019). Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. Nat Commun *10*, 4706.

Tanaka, Y., Sano, S., Nieves, E., De Libero, G., Rosa, D., Modlin, R.L., Brenner, M.B., Bloom, B.R., and Morita, C.T. (1994). Nonpeptide ligands for human gamma delta T cells. Proc Natl Acad Sci U S A *91*, 8175-8179.

Tanigaki, K., Tsuji, M., Yamamoto, N., Han, H., Tsukada, J., Inoue, H., Kubo, M., and Honjo, T. (2004). Regulation of alphabeta/gammadelta T cell lineage commitment and peripheral T cell responses by Notch/RBP-J signaling. Immunity *20*, 611-622.

Vantourout, P., Mookerjee-Basu, J., Rolland, C., Pont, F., Martin, H., Davrinche, C., Martinez, L.O., Perret, B., Collet, X., Perigaud, C.*, et al.* (2009). Specific requirements for Vgamma9Vdelta2 T cell stimulation by a natural adenylated phosphoantigen. J Immunol *183*, 3848-3857.

Vermijlen, D., Ellis, P., Langford, C., Klein, A., Engel, R., Willimann, K., Jomaa, H., Hayday, A.C., and Eberl, M. (2007). Distinct cytokine-driven responses of activated blood gammadelta T cells: insights into unconventional T cell pleiotropy. J Immunol *178*, 4304-4314.

Wilhelm, M., Kunzmann, V., Eckstein, S., Reimer, P., Weissinger, F., Ruediger, T., and Tony, H.P. (2003). Gammadelta T cells for immune therapy of patients with lymphoid malignancies. Blood *102*, 200-206.

Yin, S., Mao, Y., Li, X., Yue, C., Zhou, C., Huang, L., Mo, W., Liang, D., Zhang, J., He, W.*, et al.* (2015). Hyperactivation and in situ recruitment of inflammatory Vdelta2 T cells contributes to disease pathogenesis in systemic lupus erythematosus. Sci Rep *5*, 14432.

Yuan, J.S., Kousis, P.C., Suliman, S., Visan, I., and Guidos, C.J. (2010). Functions of notch signaling in the immune system: consensus and controversies. Annu Rev Immunol *28*, 343-365.

Zhao, Y., Niu, C., and Cui, J. (2018). Gamma-delta (gammadelta) T cells: friend or foe in cancer development? J Transl Med *16*, 3.

Zhou, J., Kang, N., Cui, L., Ba, D., and He, W. (2012). Anti-gammadelta TCR antibody-expanded gammadelta T cells: a better choice for the adoptive immunotherapy of lymphoid malignancies. Cell Mol Immunol *9*, 34-44.

List of Publications

1. Indumathi Patta*, **Ayush Madhok***, Satyajeet Khare, Kamalvishnu P Gottimukkala, Shilpi Giri, Vishal Dandewad, Jyoti Misra-Sen, Vasudevan Seshadri, Girdhari Lal, and Sanjeev Galande. Dynamic regulation of chromatin organizer SATB1 via TCR induced alternative promoter switch during T-cell development. *Nucleic Acid Res.* (2020). 48:5873-5890. PMID: 32392347 (*Equal contribution)

2. **Ayush Madhok**, Bhat Sajad, Philip Chinna, Suresh Shalini, Chiplunkar, Shubhada and Galande, Sanjeev. Transcriptome Signature of Vγ9Vδ2 T Cells Treated with Phosphoantigens and Notch Inhibitor Reveals Interplay Between TCR and Notch Signaling Pathways. *Frontiers in Immunology*. (2021) 12:660361. doi: 10.3389/fimmu.2021.660361. eCollection 2021. PMID: 34526984

3. SP Khare*, **Ayush Madhok***, I Patta, et al. Differential expression of genes influencing mitotic processes in cord blood mononuclear cells after a pre-conceptional micronutrient-based randomised controlled trial: Pune Rural Intervention in Young Adolescents (PRIYA). *Journal of Developmental Origins of Health and Disease*. 2023;14(3):437-448. doi:10.1017/S204017442200068X (*Equal contribution)

4. **Ayush Madhok**, Anjali DeSouza, and Sanjeev Galande, Understanding Immune System Development: An Epigenetic Perspective. *Epigenetics of the Immune System, Volume 17, Translational Epigenetics,* (pp 39-76). Edited by Prof. Dieter Kabelitz and Dr. Jaydeep Bhat, Published by Academic Press/Elsevier, Inc (2020).

5. Dandia, H.Y., Pillai, M.M., Sharma, D. Suvarna, M., Dalal, N., **Madhok, A**., Ingle, A., Chiplunkar, S. V., Galande, S. and Tayalia, P. Acellular scaffold-based approach for in situ genetic engineering of host T-cells in solid tumor immunotherapy. *Military Med Res* **11**, 3 (2024). https://doi.org/10.1186/s40779-023-00503-6.

6. Ajay Kumar Saw, **Ayush Madhok**, Anupam Bhattacharya, Soumyadeep Nandi, Sanjeev Galande. Combined promoter-capture Hi-C and Hi-C analysis reveals a fine-tuned regulation of 3D chromatin architecture in colorectal cancer. *bioRxiv* 2022.11.08.515643; doi: https://doi.org/10.1101/2022.11.08.515643

Manuscripts in preparation:

Ayush Madhok, Indumathi Patta, Isha Desai, Mohd Tayyab, Geeta Narlikar and Sanjeev Galande. Satb1 regulates genome conformation and gene expression along with the cohesin complex in T cell development.