FUNCTIONAL ELEMENTS AND REGULATORY INFLUENCE OF THE T CELL RECEPTOR ALPHA LOCUS CONTROL REGION

by

Janette Gomos Klein

A dissertation submitted to the Graduate Faculy in Biology.
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York
2008

UMI Number: 3325402

Copyright 2008 by
Klein, Janette Gomos

All rights reserved

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform 3325402
Copyright 2008 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

©2008

Janette Gomos Klein

All rights reserved

This manuscript has been read and accepted for the Graduate Faculty in Biology in Satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

______________ ______________________________
Date Chair of Examining Committee, Ben Ortiz

______________ ______________________________
Date Executive Officer, Laurel Eckhardt

Laurel Eckhardt

Patricia Rockwell

Erik Falck-Pedersen

Derek Sant'Angelo

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

FUNCTIONAL ELEMENTS AND REGULATORY INFLUENCE OF THE T CELL RECEPTOR ALPHA LOCUS CONTROL REGION

by

Janette Gomos Klein

Advisor: Benjamin D. Ortiz, Ph.D.

Gene regulation is an essential component in biological processes especially in complex mammalian genomes where loci have differentially regulated genes. An important part of ensuring proper gene regulation are components that activate and maintain separate domains allowing for specific gene expression profiles.

Locus control regions (LCRs) are non-coding DNA elements that regulate gene expression. LCRs provide high-level, position-independent expression to a randomly integrated transgene in a tissue specific manner. It achieves this through DNA sequences supporting multiple activities that enhance transcription and protect the expression of a transgene locus from integration site-dependent position effects. Except for elements with classical enhancer function, most key elements of the LCR activity are not well understood.

These studies used randomly integrated reporter transgene systems to identify functional elements and assess the role of the LCR within the mouse T cell receptor (TCR) alpha/Dad1 gene locus. Our investigations identified CTCF-independent regions that specifically help suppresses position effects due to random integration. Interestingly, we also found that LCR activity is orientation independent even though the endogenous gene downstream of it is ubiquitously active. In addition, our studies revealed separate 5' and 3' directed regulatory activities of the LCR. Further studies using a large, more endogenous, chromatin-like bacterial artificial chromosome (BAC) containing reporter genes flanking the LCR was also designed to further address some functions of gene regulation in this complex locus.

ACKNOWEDGMENTS

I would like to acknowledge all the members of the Ortiz Lab. Thank you especially to all the students that I have had the honor to teach and work with, namely Blanca Andino, Karl Erhard, Miguel Edwards and Randy Arroyave. They made science a new adventure every day.

Dr. Ben Ortiz continues to be one of the best teachers I have experienced both in the classroom and in the laboratory. Thank you for your mentorship, support, and encouragement. Your dedication to the laboratory and the students at Hunter College has made a truly made difference and inspired many to pursue scientific careers.

I would also like to thank Dr. Adrienne Alaie for giving me the great opportunity to teach some of her students. Her encouragement and insight always brought a smile to my face.

Thank you to the Hunter College animal facility for their help and excellent care of our animals. Thank you, Joon Kim, for your flow cytometry experience and advice.

I am grateful to the Center for Gene Structure and Function for supporting me throughout my graduate career. Not only did they provide financial support, but also provided a forum to share in its collaborative spirit, fostered opportunities to take on leadership roles and were a consistent source of exposure to new technologies.

It has been an honor to have Laurel Eckhardt, Patricia Rockwell, Derek Sant'Angelo, and Erik Falck-Pedersen as my dissertation advisors. Thank you for sharing your candid advice and scientific insight throughout the years.

Lastly, I would like to extend the greatest of appreciation to my family, past and present. Without their dedication, love and encouragement, I would not have had the opportunity or support to pursue higher education. Thank you Dad. Thank you Randy, JP, Erin and Leslie Ann. I love you.

TABLE OF CONTENTS

COPYRIGHT ii

APPROVAL iii

ABSTRACT iv

ACKNOWLEGEMENTS vi

TABLE OF CONTENTS viii

LIST OF FIGURES xi

CHAPTER 1: INTRODUCTION

1.1 T cell development 1

1.2 T cell receptors (TCRs) 4

1.3 Gene Regulation: cis-regulatory elements 6

1.4 Locus control regions (LCRs) 7

1.5 Proposed models of LCR mechanism 11

1.6 TCRα/TCRδ/Dad1 locus 13

1.7 TCRα LCR 16

1.8 Significance 23

CHAPTER 2: MATERIALS AND METHODS

2.1 DNA constructs 24

2.2 Transgenic mice 27

2.3 RNA extraction 28

2.4 Northern blot analyses 28

2.5 BioInformatic Sequence Analysis 29

2.6 Electromobility Shift Assay (EMSA) 29

2.7 Flow cytometry 30

OVERVIEW: Chapters 3-5 31

CHAPTER 3: CTCF-independent elements significantly contribute to TCRα locus control region activity

3.1 RESULTS
3.1.1 Candidate regions for TCRα LCR functional studies 33
3.1.2 Transgenic mouse models 34
3.1.3 Transgene expression levels due to LCR deletion mutations 37
3.1.4 Elements that contribute to position effect suppression 42

3.2 SIGNIFICANCE
3.2.1 Discrete functional sub-regions of the TCRα LCR identified 46
3.2.2 Insulator-like activity of HS4 and HS6-316 46

3.3 FUTURE DIRECTION
3.3.1 Conserved Sequences in HS4 and HS6-316 51
3.3.2 Binding sites in HS4 and HS6-316 53
3.3.3 In vivo footprint 55
3.3.4 Isolate sequences conferring insulation capacity 56

CHAPTER 4: Regulatory Influence of the TCRα LCR

4.1 RESULTS
4.1.1 Transgenic mouse models 58
4.1.2 Transgene expression levels are related to copy number regardless of TCRα LCR orientation 60
4.1.3 Transgenes flanking either side of the TCRα LCR display 61

TCRα- like expression
4.1.4 hCD2 ectopic B cell expression is suppressed in CD2:B7 transgenes 64
4.1.5 HLAB7 Ectopic B cell expression persists in transgenes placed 3' of the TCRα LCR 68
4.1.6 TCRα LCR cannot maintain hCD2 Peripheral T cell expression 69

4.2 SIGNIFICANCE
4.2.1 TCRα LCR activity is orientation independent 71
4.2.2 LCR activity is influenced differently when flanked by separate reporter genes 72

4.3 FUTURE DIRECTION
4.3.1 DNase1 hypersensitivity assays (DHA) 75
4.3.2 Chromatin Immunoprecipitation (ChIP) 76

CHAPTER 5: Creating a BAC reporter transgene model to study TCRα/Dad1 regulation

5.1 RESULTS
5.1.1 Bacterial Artificial Chromosomes (BAC) 78
5.1.2 BAC containing the TCRα LCR 78
5.1.3 Modified TCRα LCR BAC 79
5.1.4 Modified BAC sequence 81
5.1.4 Isolating BAC fragment for transgenic injection 81

5.2 SIGNIFICANCE
5.2.1 TCRα reporter gene expression 82
5.2.2 Dad1 reporter gene expression 83

5.3 FUTURE DIRECTION
5.3.1 Copy number of transgenic lines 84
5.3.2 Expression of reporter transgenes 84
5.3.3 BAC model system 85

CHAPTER 6: Discussion 86

REFERENCES 89

LIST OF FIGURES

Figure 1. Overview of T cell development from the hematopoietic stem cell (HSC). 2

Figure 2. Genomic locus of the TCRα/TCRδ/Dad1. 14

Figure 3. Overview of TCRα LCR characterization. 17

Figure 4. TCRα LCR and enhancer blocking regions identified within it. 22

Figure 5. TCRα LCR and β-globin transgene constructs. 36

Figure 6. Minor effect of transgene expression levels per copy due to deletion of CTCF sites in the TCRα LCR. 38

Figure 7. Transgene expression levels per copy are significantly reduced upon internal deletion of either the HS4 or HS6-316 regions. 40

Figure 8. Summary of data from multiple experiments showing the effect of internal deletions on TCRα-driven expression levels in lymphoid organs downstream of HS1. 41

Figure 9. Line-to-line consistency in relative tissue distribution of transgene expression from wild-type and mutant LCR-driven constructs. 45

Figure 10. Insulators can help define a regulatory region. 48

Figure 11. Conserved DNA sequences exist within HS4 and HS6-316. 52

Figure 12. DNA binding sites within conserved regions of HS4 and HS6-316. EMSA experiments showing formation of complexes with conserved HS4 and HS6-316 regions. 55

Figure 13. TCRα LCR and transgene constructs containing hCD2 and/or HLA-B7 reporter genes. 59

Figure 14. Transgene expression correlates to copy number regardless of orientation. 61

Figure 15. Pattern of expression in transgenes is similar to TCRα expression. 63

Figure 16. Ectopic expression of upstream TCRα LCR-linked reporter gene (hCD2) is suppressed by adding a downstream gene (HLA-B7) flanking the LCR. 65

Figure 17. Mean fluorescence intensity (MFI) of hCD2 reporter staining in paired samples of whole thymocytes and spleen B cells. 66

Figure 18. Ectopic hCD2 RNA levels are suppressed in CD2:B7 transgenic lines. 67

Figure 19. Spleen B cell expression (relative to thymus) of the HLA-B7 reporter is unchanged by the addition of a second reporter gene upstream of the LCR. 69

Figure 20. Peripheral T cell expression is not maintained in the double CD2:B7 transgenic lines. 70

Figure 21. Model of bi-directional LCR gene expression regulation 74

Figure 22. Diagram of the TCRα/Dad1 Bacterial artificial chromosome (BAC) reporter construct, Δ5'I14h/rCD2 79

Chapter 1: Introduction

1.1 T cell development

Overview:

T cells are an important component of the adaptive immune response. Adaptive immune responses direct lymphocytes in an antigen specific, cell-mediated response to eliminate pathogens. In order for T cells to be successful in the immune response process, it must be able to bind to and distinguish a foreign antigen in the context of an antigen presenting molecule. T cell receptors (TCRs) mediate this interaction. Proper regulation of the genes encoding TCR subunits is essential for T cell development. T-cells must undergo a strictly regulated series of developmental steps to reach functional maturity (Figure 1).

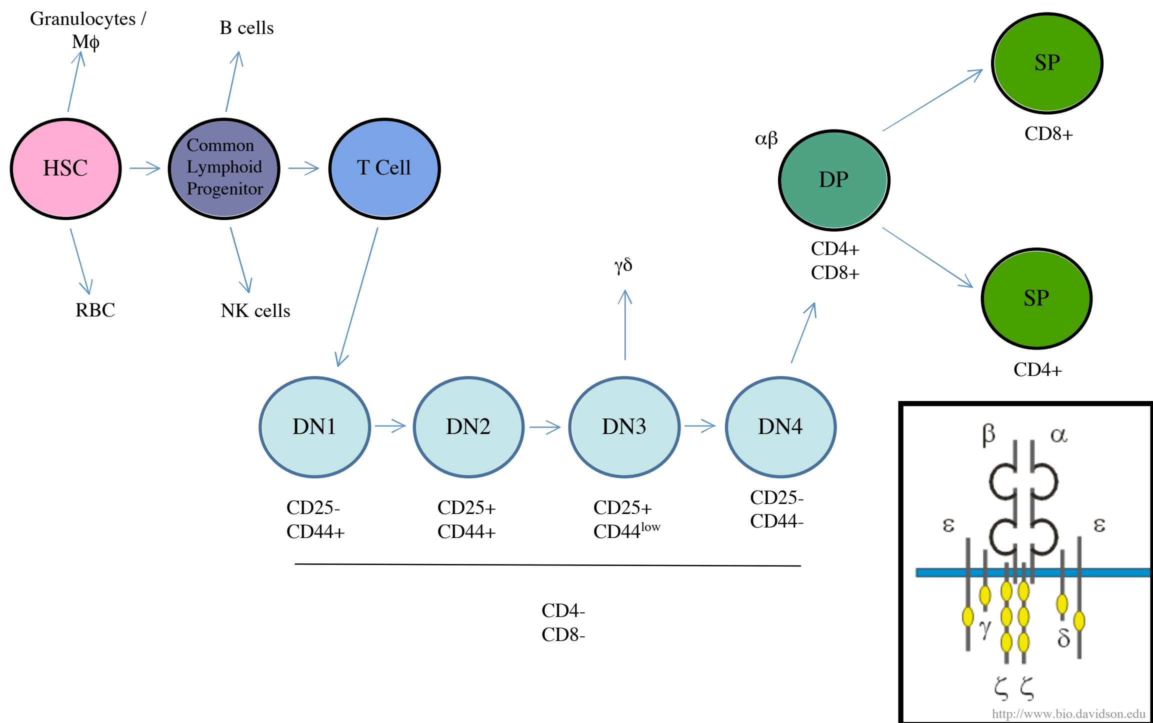


Figure 1. T cell development. Overview of T cell development from the hematopoietic stem cell (HSC). The HSC, residing in the bone marrow, is capable of differentiating into all the blood cells types in the lymphoid and myeloid lineages. Among others, circulating T cells develop from the common lymphoid progeniter which migrates to the thymus. T cell development in the thymus can be followed by monitoring the expression of a series of cell surface proteins: CD4, CD8, CD25, and CD44. The double negative (DN) T cell population does not express CD4 or CD8 and is further subdivided by CD25 and CD44 expression. Interactions between transcription factors and gene regulatory elements direct lineage decisions and progression to functional T cells. Inset: Cartoon of the αβ T cell receptor. The α and β chains form a heterodimer that can bind antigen plus major histocompatibilty complex (MHC) while CD3, made up of the ε, γ, δ, ζ peptide chains, contribute to downstream signaling.

T-cell development in the thymus:

T cells develop from pluripotent hematopoietic stem cells (HSC) that originate in the bone marrow (Ogawa 1993). HSC become progenitor cells called the common lymphoid

progenitor that continue their development in the thymus to become thymocytes [reviewed in(Schwarz and Bhandoola 2006)]. In the thymus, stages of T cell development are delineated by cell surface markers (Figure 1).

The most immature thymocytes do not express cell surface markers for CD4 or CD8 (CD4-CD8-) and are called double negative (DN)(Godfrey et al. 1993; Pearse et al. 1989). This DN stage is further subdivided into four stages based on the expression levels of CD25 and CD44: DN1 (CD25-CD44+), DN2 (CD25+,CD44+), DN3 (CD25+,CD44-), and DN4 (CD25-, CD44-). At specific points in the DN sub-stages, T-cell receptor (TCR) gene segments are productively rearranged to create unique antigen binding sites (Godfrey et al. 1994; Wilson, Held, and MacDonald 1994; Hoffman et al. 1996).

As T cell precursors continue to develop, CD4 and CD8 are expressed (double positive, DP) and functional TCRs are expressed (Godfrey et al. 1994; Wilson, Held, and MacDonald 1994; Hoffman et al. 1996; Fehling et al. 1995; Dudley et al. 1994). The TCR interacts with cortical epithelial cells that express a high density of MHC class I and class II molecules (Goldrath and Bevan 1999). Interaction of the TCR with these self-peptide–MHC ligands dictates the fate of the DP thymocytes .

Negative and positive selection occurs ensuring that unfit T-cells cannot leave the thymus [reviewed in (Baldwin, Hogquist, and Jameson 2004)]. This process eliminates T cells that do not bind self-MHC, are auto-reactive or do not provide the appropriate level of signaling. Those thymocytes that express TCRs that bind MHCI become CD8+ while T

cells expressing TCRs that bind MHCII ligands become CD4+ T cells. The single positive (SP) naïve T cells are then released and populate peripheral lymphoid organs, such as the spleen and lymph nodes.

1.2 T cell receptors (TCRs)

There are 4 TCR subunits: α, β, γ, and δ which are each encoded by independent genes. TCRs contain a heterodimer with TCR alpha (TCRα) and TCR beta (TCRβ) peptide chains in αβ T cells, or TCR gamma (TCRγ) and TCR delta (TCRδ) in γδ T cells. Greater than 95% of circulating T cells express the αβ peptide chains. While a small number of γδ T cells circulate, most become localized to the gut mucosa and epidermis (Brenner, Strominger, and Krangel 1988; Triebel and Hercend 1989).

Rearrangement:

T lymphocytes are able to recognize a vast repertoire of antigen binding sites because of their ability to rearrange the antigen-binding portion of their receptor [reviewed in (Sleckman 2005)]. Among the enzymes essential to this process are recombinase activating genes, RAG1 and RAG2 (Wilson, Held, and MacDonald 1994).

TCRs rearrange their gene segments similarly to the immunoglobulin heavy chain. To make a functional TCR, the TCR genes first assemble variable (V), diversity (D) and

joining (J) or VJ gene segments. Rearrangement of TCR β, γ, and δ gene segments occurs before TCRα at early immature DN2 and DN3 T-cell stages (Wilson and Stauffer 1994; Godfrey et al. 1994). TCRα genes rearrange at later DN4 and DP stages (Lauzurica and Krangel 1994; Hernandez-Munain, Sleckman, and Krangel 1999; Krangel et al. 2004).

Cell Surface Expression:

Productively rearranged TCR gene products are assembled into a receptor complex. When TCRβ is assembled, a surrogate pre-TCRα stabilizes the complex until a functional TCRα chain is made(Saint-Ruf et al. 1994; Groettrup et al. 1993). TCRδ and γ rearrange at the same time. Although the exact mechanism of commitment to the γδ lineage is not known, it is thought that if TCRγ and δ first are able to make a productive rearrangement, TCRβ rearrangement stops and γδ T cells are produced (Kang and Raulet 1997; Pardoll et al. 1987).

TCR complex and accessory membrane molecules:

In addition to heterodimeric chains of αβ or γδ TCR subunits, several other units are required for full TCR function. The CD3 complex associates with the TCR and consists of five polypeptide chains that associate to form three dimers. CD3 is required for membrane expression of the complex and participate in downstream signal transduction after antigen-MHC binding (Punt et al. 1994; Malissen et al. 1999). In addition, the co-

receptors CD4 or CD8 bind to conserved regions of MHC (either Class I or class II). Their binding serves to transmit stimulatory signals to T cells [reviewed in (Vukmanovic and Santori 2005)]. Other accessory cell-adhesion molecules strengthen the interaction between a T cell and its target cell. This

1.3 Gene Regulation: cis-regulatory elements

The controls that act on gene expression are complex and regulated at many levels. Transcriptional activation of genes requires a structured recruitment of basal transcription machinery that works in concert with other elements that can make changes in chromatin structure. Traditionally, studies looking at cis-elements that effect eukaryotic gene expression were limited to gene-proximal elements such as promoters, enhancers, and silencers. Promoter or enhancer sequences allow for increase of gene expression by influencing the activity and direct binding of RNA polymerase. These cis-regulatory elements are part of the larger transcriptional machinery that controls the production of gene products in a specific manner.

The identification of other cis-regulatory elements that influence gene expression has increased dramatically beyond promoters and enhancers. Examples are unmethylated CpG islands, insulator binding sites, scaffold/matrix attachment regions (S/MARs), and locus control regions (LCRs). These elements affect chromatin structure, interaction, and nuclear localization. In addition, these cis-regulatory elements affect the ability for

factors to form complexes or access specific genes. Thus, developmental and cell lineage-specific regulation of gene expression relies on higher order interactions including long-range interactions of various cis-regulatory elements and dynamic chromatin alteration.

Most recently, analysis of the spatial organization of chromosomes reveals complex three-dimensional networks of chromosomal interactions. Gene expression can be effected in multiple ways, including interactions by long-range control by locus control regions, enhancers and repressors, coordinated expression of genes, and modification of epigenetic states. Understanding the functional consequences of these associations is an important endeavor in the effort to understand the dynamic relationships of gene regulation.

1.4 Locus control regions (LCRs):

Characteristics:

LCRs are composed of a group of cis- DNA regulatory elements that lie within DNase I hypersensitivity (HS) sites. Hypersensitivity occurs because of the lack of the highly compact nucleosome structure. These sites indicate where regions of DNA are in a selectively accessible state, and thus are assumed to be “active” and important in transcriptional regulation. The core determinants at individual HSs are usually composed of multiple ubiquitous and/or lineage- specific transcription factor binding sites that carry

out specific functions. These include classical enhancer and, insulator-like activity, and others whose functions are essential for LCR activity, but whose mechanism of action is still unknown. Often, LCRs are remote from the genes that they regulate.

While their composition and location relative to their cognate genes can differ, they share the common property of maintaining physiological levels of gene expression. Locus control regions are described as cis- DNA elements that maintain a functional domain in an open, active configuration in chromatin. Although the mechanisms of LCR function are not well understood, it has been shown that LCRs are required in some loci for high level and tissue/stage specific expression. Locus control regions are functionally defined by their ability to enhance the expression of linked genes to physiological levels in a tissue-specific and copy number-dependent manner at ectopic chromatin sites. A transgene linked to an LCR can suppress the regulatory effects of adjacent heterochromatin and thus become protected against the phenomenon of position effect variegation or silencing. In this way, transgenic expression in mice correlates linearly with transgene copy number. In summary, LCRs are sequences that contain DNA sequence information sufficient for establishing full, position-independent expression of a transgene. This suggests a role for LCRs in selectively opening up extensive chromatin domains.

History:

The first LCR was identified over 20 years ago. While examining β-globin gene sequences in patients suffering from Dutch γδ-thalassemia, it was found that the gene structure was normal but a large deletion of upstream sequences was associated with abnormal β-globin expression and disease (Kioussis et al. 1983). In transgenic mouse studies, it was determined that a LCR was required for normal regulation of β-globin gene expression (Grosveld et al. 1987). Much of what we know about LCR function has come from the study of the chicken, human, and mouse β-globin LCR.

LCRs in immunologically important loci:

Since their discovery, data from transgenic model systems indicate that LCRs are needed for the regulation of chromatin structure. Numerous locus control regions have been identified in mammalian gene systems. Many of these LCRs have been identified in gene loci containing immunologically important genes such as CD2, TH2 cytokines locus, major histocompatibility complex class II (MHCII) locus and TCRα/δ.

CD2 is a cell-surface receptor molecule expressed on nearly all thymocytes. Among CD2 functions are mediation of T cell adhesion and participation in signal transduction via its cytoplasmic domain (Moingeon et al. 1989; Chang et al. 1989). The hCD2 LCR was shown to be essential for establishing an open chromatin configuration, even in the absence of enhancer activity (Festenstein et al. 1996; Greaves et al. 1989).

The TH2 LCR is at the 3' end of a ubiquitously expressed gene (Rad50) that is flanked by IL5 and IL13. Intra-chromosomal looping occurs that allows the TH2 LCR to associate with TH2 cytokine gene promoters (IL4, IL5, IL13) upon naïve CD4+ T cell differentiation (Spilianakis and Flavell 2004; Lee, Spilianakis, and Flavell 2005). Interestingly, this TH2 LCR participates in inter-chromosomal interactions. In these studies, the TH1 Ifnγ promoter (chromosome 7) interacts with the promoters of IL5 and Rad 50 (chromosome 11) as well as the TH2 LCR in CD4+ cells that have not yet differentiated down the TH1 or TH2 pathway (Spilianakis et al. 2005).

Upstream of the human DRα (HLA-DRA), and orthologous murine H-2 Eα promoter, is the MHCII LCR. Within this LCR are binding sites called Y'-S' that interact with regulatory factor X complex (RFX) and co-activator CIITA (Krawczyk et al. 2004). The promoter of HLA-DRA and H-2 Eα has inverted matching binding motifs called the S-Y region (Ting and Trowsdale 2002). It is possible that these essential activators of MHC class II transcription dimerize to form a loop between the LCR with its promoter. In addition, the binding for RFX and CIITA to the LCR induces bidirectional long-range histone acetylation (Masternak et al. 2003).

The TCRα LCR lies within a complex locus containing lymphoid specific genes (TCRα and TCRδ) as well as a ubiquitously expressed (Dad1) gene. TCR expression is essential for proper T cell development. Among the LCR's regulatory units, tissue specific methylation (Santoso, Ortiz, and Winoto 2000) and insulator-like activity has been observed (Magdinier, Yusufzai, and Felsenfeld 2004; Zhong and Krangel 1999; Ortiz et

al. 1997; Ortiz et al. 2001). This locus and its LCR are the focus of my thesis and will be discussed later in greater detail.

Numerous LCRs have been identified in mammals. Their properties are influential both in its endogenous locus and in transgenic studies. Thus, LCRs are potent activators and regulators for transcription and appear to operate by ensuring an open chromatin configuration. However, the mechanisms of LCR activity are still unknown.

1.5 Proposed models of LCR mechanism

Although the mechanisms of LCR activity are unknown, several models exist for how the LCR exerts its effects on chromatin and transcription from a distance. These models include looping, tracking, linking, and the active chromatin hub (Li et al. 2002; Harju, McQueen, and Peterson 2002; Engel and Tanimoto 2000; de Laat and Grosveld 2003).

Looping:

In this model, transcription factors bind the LCR and gene promoter. The LCR directly interacts with the promoter causing a "looping out" of any intervening DNA. The interactions form an active transcription holocomplex allowing for transcriptional competence. Here, the LCR section of the DNA loops around to effect transcription so that the binding site of the transcriptional machinery (proteins) is in physical close

proximity to bridge and affect the transcription of that specific gene. Evidence of these models exists in the β-globin locus and TH2 cytokines locus (Carter et al. 2002; Tolhuis et al. 2002; Spilianakis and Flavell 2004).

Tracking/Facilitated tracking:

In the tracking model, transcription factors bind the LCR sequence to form an activation complex that migrates linearly or “tracks” along the DNA locus. Once a promoter is found, a stable structure is established. The transcriptional apparatus can assemble and transcription is initiated. Supporting this model, intergenic nuclear transcripts has been detected across some LCRs (Blackwood and Kadonaga 1998; Tuan, Kong, and Hu 1992).

Linking:

The linking model suggests that propagating factors bind DNA to define a transcriptional domain. A promoter alone is unable to recruit the crucial activity that is provided after initiation of a chain of these factor-containing complexes by an LCR. The linking model proposes there is a sequential stage-specific binding of transcription factors and chromatin facilitator proteins throughout the locus. In loci where regulatory elements are dispersed across large regions, each distinct element serves to reinforce an overall structure that is propagated throughout the locus. Supporting this is a widely expressed chromosomal protein in Drosophila called chip. Chip- containing complexes act by binding chromatin and organizing intervening sequences that allow enhancer-promoter

interactions to occur while the LCR serves to open the chromatin domain (Bulger and Groudine 1999; Dorsett 1999).

Active Chromatin Hub (ACH):

In the ACH model, the LCR initially serves as a multiple element receptor that acts as a hub for factor binding to direct chromatin remodeling (de Laat and Grosveld 2003). This ACH can then act as an enzyme. Once chromatin-remodeling activity is complete, the LCR directly interacts with downstream genes, facilitating a proximity-dependent activation of gene expression by the LCR. Evidence in the β-globin locus show that active β-globin genes physically interact in the nuclear space with multiple cis-regulatory elements while inactive genes are looped out. Deletion of the 5'HS2, HS3, or HS3 core regions of the β-globin LCR has provided a potential mechanistic evidence for this model (Patrinos et al. 2004).

1.6 TCRα/TCRδ/Dad1 locus

Size and Orientation:

Complex loci are likely to provide insight on higher orders of gene regulation. The T-cell antigen receptors, alpha and delta subunits / Defender against cell death 1 (TCRα/TCRδ/Dad1) locus is large and complex. It spans over 1.5Mb and is highly

conserved (Koop et al. 1994; Wang et al. 1997). It is located on chromosome 14, specifically at location 14q11.2 in human and 14 19.5 cM in the mouse. The majority of the 5' portion in this locus is subject to gene segment rearrangement. In addition, this locus contains 3 differentially expressed genes: TCRα, TCRδ, and Dad1. Dad1 lies in the opposite transcriptional orientation as TCRα and TCRδ. There are several regulatory elements within the locus including the TCRα LCR that lies between the TCRα and Dad1 genes. The layout of the TCRα locus is shown in Figure 2.

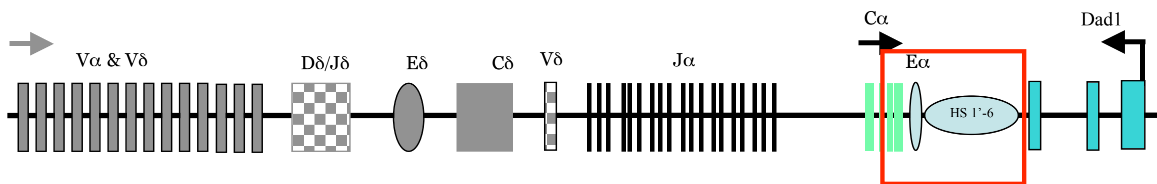


Figure 2. Genomic locus of the TCRα/TCRδ/Dad1 genes (not drawn to scale). The locus is over 1.5 Mb consisting mostly of V and J gene segments. The borderless light green boxes represent the exons of the TCRα constant region. The teal colored boxes indicate the Dad1 exons. The red box indicates the TCRα LCR region. The arrows indicate transcriptional orientation.

TCRα/TCRδ:

Although similar, these genes are expressed on separate cell types and during different stages of T cell development. While rearrangement of TCRδ occurs at early immature DN2-DN3 stages, TCRα genes rearrange at DN4 and DP stages. Contained completely within the germline TCRα gene is the TCRδ gene. The TCRδ gene resides within the Vα and Jα gene segments of TCRα. Thus Vα and Jα joining results in the entire TCRδ gene

being deleted from the genome. Thus, only TCRα or TCRδ can become expressed on an αβ or γδ T cell respectively.

Dad1:

Also part of the TCRα locus is the Dad1 gene that is ubiquitously expressed and located 12kb 3' of the TCRα constant region (Cα). Dad1 is expressed in both non-lymphoid and lymphoid organs and can be detected as early as embryonic day seven, well before TCR gene expression (Hong et al. 1997). Dad1 is an anti-apoptotic gene that encodes a subunit of an oligosaccharyltransferase complex that initiates N-linked glycosylation (Nakashima et al. 1993; Kelleher and Gilmore 1997). The role of Dad1 in T cell development is not understood but Dad1 expression is modulated during T cell development (Hong et al. 1999) and a Dad1 knock-out is embryonic lethal (Hong et al. 2000). Dad1 expression rises as T cells enter the single positive stage correlating with an increase in TCRα expression (Hong et al. 1999).

TCRα LCR:

How is this complex locus, supporting the differential gene expression patterns of three separate genes, regulated? Among the regulatory elements identified to date includes a locus control region, the TCRα LCR. The TCRα LCR spans an 11.5kb region that lies between the TCRα gene and Dad1. Studies in this and other LCRs suggest that this

regulatory unit may play a role in the regulation of multiple aspects of this locus. There are numerous reports of the TCRα LCR regulating TCRα gene expression. In addition, Dad1 expression is upregulated during T cell development where LCR activity is seen. Furthermore, several enhancer blocking and barrier-like insulator regions have been identified within the LCR. The TCRα LCR lies in close proximity to both TCRα and Dad1. These aspects of the LCR provide support for hypothesizing its involvement in multigene regulation. A more detailed description of the TCRα LCR will follow.

1.7 TCRα LCR

The LCR resides over a large region that extends from within the TCRα chain constant region to 3kb upstream of the 3' end of Dad1 gene. In its endogenous locus, it measures about 11.5kb. It is characterized by nine DNase 1 hypersensitive sites (HS) that synergize in a way that is not completely understood. The 9 defined DNase1 HS sites of the TCRα LCR are called 7-8, 1, 1', 2-6. Of these HS sites, 1, 1', 4 and 6 have the strongest HS. This region that defines the LCR can direct physiological, position-independent, copy number dependent, and lymphoid (T cell) specific expression of a linked transgene (Diaz, Cado, and Winoto 1994; Ortiz et al. 1997). Later studies linked the TCRα LCR to a reporter gene that allowed analyses of various cell types and during stages of T-cell development (Harrow and Ortiz 2005). It was found that the TCRα LCR was able to specify upregulation of thymic expression at the DN4 thymocyte stage and maintain high levels throughout subsequent thymic developmental stages. Following thymic

development, endogenous TCRα normally gets upregulated in the peripheral T cells. However, LCR linked reporter gene expression in the periphery was low indicating that additional gene regulatory elements are needed for this aspect of TCRα expression. Other data investigating specific hypersensitive sites of the TCRα LCR, mostly through transgenic studies, has revealed some of their functions that will be described in detail below (summarized in Figure 3).

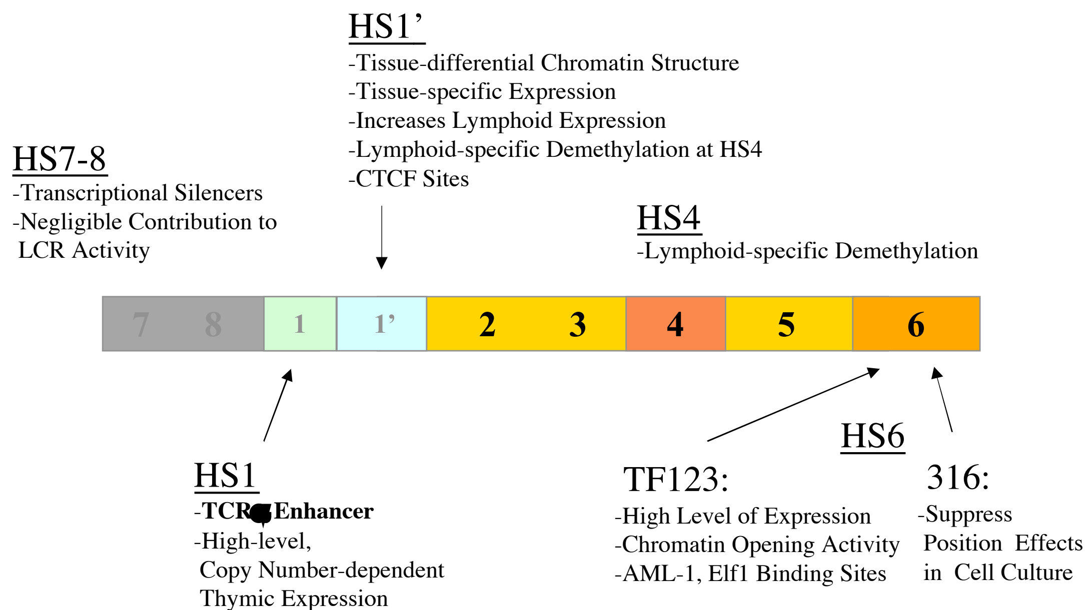


Figure 3. Overview of TCRα LCR characterization. Several regions have been studied for LCR function. This is a brief summary of LCR activity found thus far. Number indicates hypersensitivity site. The HS in grey indicate the 5' tissue specificity region while the black numbers represent the 3' widely active "chromatin-opening" region.

Sub-elements of the LCR:

The TCRα LCR can be divided into a 5' region containing hypersensitive sites (HS) 7-8, 1-1' and a 3' region containing HS 2-6. In addition more discrete regions of the LCR have been analyzed and are described below.

The 5' region of the TCRα LCR:

The 5' region, containing HS7-8, 1,1', has been designated as the region that influences tissue specificity. The 5' region includes previously characterized transcriptional silencers in the region of HS 7-8 (Winoto and Baltimore 1989). The TCRα enhancer (Eα) is located in HS1 (Winoto and Baltimore 1989). A deletion of the HS1/1' region of the endogenous TCRα LCR significantly impairs the function of the TCRα gene locus (Sleckman et al. 1997). A similar deletion affected high level, copy- number dependent, tissue specific expression and differential chromatin structure in transgene assays (Ortiz, Cado, and Winoto 1999; Ortiz et al. 1997). However, a deletion of only HS1 did not change tissue specificity or chromatin structure in different tissues.

HS7-8:

HS7 lies within the TCR Cα introns. HS7-8 contains transcriptional silencers that were defined in transient transfection assays (Winoto and Baltimore 1989). The HS7-8 region silenced Eα enhancer activity in several types of cells but not αβ T cells. HS7-8 has been deleted in constructs containing the LCR with a linked transgene. Because HS1-6 seems

to have full LCR activity in LCR-linked transgenic studies, a role for HS7-8 in vivo has not yet been supported.

HS1:

HS1 contains the Eα enhancer and its HS is strongly active in normal thymocytes. Eα is the only element of the LCR that support classical transcriptional enhancer activity (Winoto and Baltimore 1989). However, this enhancer alone was unable to give position-independent and tissue-specific expression in mice. HS1 does contribute to high-level expression in thymus and contributes to copy number dependent expression (Ortiz, Cado, and Winoto 1999). This was evidenced by comparing partial linked LCR-transgenes in a transgenic construct containing HS1-6 versus HS1'-HS6.

HS1':

HS1' was shown to be a separate essential element for the TCRα LCR (Ortiz, Cado, and Winoto 1999). Two roles for HS1' has been suggested. Its presence increases expression in lymphoid tissues while suppressing ectopic transgene activity in non-lymphoid cells. No classical enhancer function has been described in HS1'. It was found that HS1' is present in all tissues examined (Hong et al. 1997; Ortiz, Cado, and Winoto 1999). Other studies on HS1' indicates that it plays an additional role in chromatin structure by affecting methylation status at other regions of the LCR (Santoso, Ortiz, and Winoto 2000). A deletion in HS1 does not contribute significantly to maintaining chromatin

structure while a deletion in HS1' causes changes in chromatin structure as seen by hypersensitivity assays and prevents the normal de-methylated status of the HS4 region in lymphoid tissues.

In addition, the HS1' region contains transcription factor CTCF binding sites that support enhancer blocking activity (Magdinier, Yusufzai, and Felsenfeld 2004; Gomos-Klein et al. 2007). CTCF is a multivalent protein that binds to different targets through its zinc finger domains and can be involved in gene activation, repression, and enhancer blocking insulation (Ohlsson, Renkawitz, and Lobanenkov 2001). It is the only known factor to support enhancer blocking insulator activity in vertebrates.

Because HS1' forms in the chromatin of non-lymphoid tissues and has the ability to affect chromatin structure, this sub-element may also play a role in separating the regulation of lymphoid specific TCRα and ubiquitously expressed Dad1 genes.

3' region- HS2-6:

The 3' region of the TCRα (HS2-6) LCR drives expression of a linked transgene in a non-tissue-specific manner. The 3' region of the TCRα LCR by itself causes expression in a variety of tissues while the entire LCR is tissue specific. Therefore, HS2-6 is considered a widely active "chromatin-opening" region. HS2-6 required for complete LCR activity. Putting HS2-6 in a linked transgene causes LCR like activity but with no apparent tissue restrictions. In these HS2-6 linked transgenics, the HSs seem to be

equally prominent in lymphoid/non-lymphoid tissues. HS2-6 also contains enhancer blocking activity (Zhong and Krangel 1999; Magdinier, Yusufzai, and Felsenfeld 2004). Thus, HS2-6 seems to counteract position effects. Its widespread activity may suggest a role for it in ubiquitous Dad1 expression.

HS4:

HS4 has been shown to be demethylated in thymus and splenic tissues and methylated in non-lymphoid tissues (Santoso, Ortiz, and Winoto 2000). This methylation pattern is established early in development (before recombination occurs) and is influenced by HS1'. Transgenes linked to HS1-8, HS 1-6, and HS1'-6 maintain HS4 demethylation at thymus and spleen. However, when a transgene is linked HS2-6, HS4 is not demethylated in thymus and spleen (Santoso, Ortiz, and Winoto 2000).

HS6:

HS6 is strongly detected in chromatin of normal thymocytes. Normally, in the context of the full LCR, HS6 appears to be active only in lymphoid organs and bind tissue restrictive factors, AML-1/Runx1 and Elf-1 (Ortiz et al. 1997; Ortiz, Cado, and Winoto 1999; Ortiz et al. 2001). The transcriptional activity of transgenes containing only HS6 is similar to transgenes containing HS2-6. Thus, HS6 is thought to contain the bulk of the activity that drives widely active integration site-independent transgene expression (Ortiz et al. 2001). In later studies, deletions within HS6 negatively affect the LCR's position

effect suppression capacity indicating its role in regulating long-range chromatin structure (Harrow et al. 2004; Gomos-Klein et al. 2007).

Insulator Enhancer-blocking regions within TCRα LCR:

The TCRα LCR contains several broad regions which demonstrate enhancer- blocking activity (Figure 4). This includes regions that are CTCF-dependent and others that are CTCF-independent (Magdinier, Yusufzai, and Felsenfeld 2004). CTCF is the only identified transcription factor in vertebrates that participates in enhancer- blocking insulation. The CTCF sites are located in HS1' (Magdinier, Yusufzai, and Felsenfeld 2004; Gomos-Klein et al. 2007). The HS2–6 region also displays enhancer-blocking activity (Zhong and Krangel 1999; Magdinier, Yusufzai, and Felsenfeld 2004). However, the precise location of these activities is still unknown.

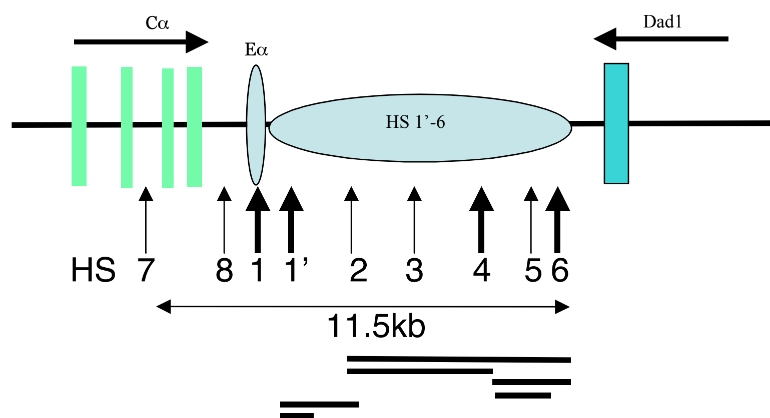


Figure 4. TCRα LCR and enhancer blocking regions identified within it. The TCRα LCR, indicated by arrowed numbers, TCRα constant region and last exon of Dad1. Horizontal arrows indicate transcriptional orientation. Hypersensitivity sites (HS) are indicated by number. Enhancer blocking regions that have been identified are indicated by thick horizontal lines (Zhong and Krangel 1999; Magdinier, Yusufzai, and Felsenfeld 2004). Only the HS1' has been shown to contain CTCF binding sites.

1.8 Significance

Elucidating the mechanisms of transcriptional control in highly regulated genes is important for therapeutic purposes. Current gene therapy applications are cursed with several problems that include non-physiological expression, gene silencing due to position effect and ectopic expression. Because LCRs are able to overcome many of the problems seen in current gene therapy applications, understanding LCR mechanisms and the elements that define it are of great importance. The inclusion of essential LCR components to gene therapy vector design may specifically overcome the regulatory issues that current vectors have.

A complex mosaic of regulatory elements regulates the temporal and spatial patterns of genes. Large loci containing differentially regulated genes such as the TCRα/TCRδ/Dad1 locus provides an ideal platform for dissecting these mechanisms.

Chapter 2: Materials and Methods

2.1 DNA constructs

β-globin reporter gene constructs:

A 4.9kb human β-globin gene fragment was linked to the 7.4kb fragment containing HS1-6 of the TCRα LCR (β:1-6) as previously described (Ortiz, Cado, and Winoto 1999). TCRα LCR deletion mutations in the β:1-6 constructs have also been described (Gomos-Klein et al. 2007). Briefly, β:1-6ΔCTCF is a deletion mutant version of β:1-6 in which a 94bp AvrII to ScaI fragment within HS1' containing tandem CTCF binding sites was removed. Similarly, the β:1-6ΔHS4 construct was designed to delete an ApaLI to DraI fragment containing 1.1 kb of HS4 region DNA from the β:1-6 wild type LCR-linked transgene. β:1-6Δ316 is a construct that removes a 316-bp PstI to BglII fragment in the HS6 region from the β:1-6 transgene. Transgene inserts for microinjection were isolated from vector DNA by using the XhoI and ClaI sites of the parent cloning vector, pSP72 (Promega, Madison WI).

hCD2 and HLA-B7 reporter gene constructs:

Previously described reporter genes, hCD2ΔT (Melton and Stoneking 1996) and/or HLA-B7 (Chamberlain et al. 1988), were linked to the TCRα LCR HS1-8 regions. The hCD2ΔT reporter gene has an introduced stop codon in exon V of the gene with the

second and third introns removed (Melton and Stoneking 1996). The HLA-B7 transgene construct was previously described and contains the genomic fragment of the human HLA-B7 gene (Chamberlain et al. 1988). HLA-B7 encodes the heavy chain of a human major histocompatibility complex class I. Neither of these fragments contain the endogenous LCR activity present in their natural loci.

The 9kb fragment containing the TCRα LCR HS1-8 regions was linked to these reporter genes (Ortiz et al. 1997). Here, the hCD2:1-8 (hCD2) construct contains the 10.5kb SalI-BamHI hCD2 fragment described above linked to the 5' end of this TCRα LCR HS1-8 fragment (Harrow and Ortiz 2005). For the TCRα LCR linked HLA-B7 (B7) construct, a 5.5kb EcoRI-BamHI HLA-B7 fragment was linked to the 3' side of the TCRα LCR HS1-8. A construct called CD2:B7 contained both of the described reporter genes, hCD2ΔT and HLA-B7. To make CD2:B7, hCD2ΔT (SalI-NruI) was inserted into the HLA-B7 construct. In this double transgene construct, hCD2ΔT was placed in the 5' orientation relative to TCRα LCR HS1-8 while HLA-B7 was at the 3' end in the endogenous Dad1 orientation. CD2:B7 transgene inserts for microinjection were isolated from vector DNA.

Modified reporter Bacterial Artificial Chromosome:

The BAC clone RP23-94I14 (I14) from the BACPAC Resources Center (BPRC) contains 46105796-46287953 base pairs from mouse Chromosome 14. This clone contains most of the TCRα /TCRδ/Dad1 locus including the TCRδ gene, TCRα Jα segments, TCRα constant region, *Dad1*, and several genes 3' of Dad1. BAC modification to delete or

insert was done by recombination steps using the Red/ET Recombination system (Genebridges, Dresden, Germany). The 182157bp I14 construct was modified to remove the region upstream of the TCRα constant region containing TCRδ and multiple Jα regions creating a 91816bp BAC construct. A separate DNA fragment, Vα11.1hCD2ΔT, containing the Vα11.1 promoter from the pVCαK transgene (Diaz, Cado, and Winoto 1994) was inserted in frame with a reporter gene, hCD2ΔT cDNA (Melton et al. 1996). The 1680bp Vα11.1hCD2ΔT construct was put in the TCRα orientation just 180bp upstream of the TCRα constant region of the BAC. Homology arms for recombination into the BAC were included in the Vα11.1hCD2 construct. Homology arms were made from the EcoRI-BglII fragment from pVCαK and from the BglII site upstream of the TCRα constant region. Similarly, a 703bp rat reporter cDNA CD2 that also lacked the cytoplasmic tail reporter gene (Hozumi et al. 2000) was inserted at the ATG of the first exon of Dad1. The rCD2 reporter gene in the Dad 1 orientation was amplified by PCR linking homology arms that matched 50bp upstream and downstream of the Dad1 ATG start site in the BAC. The resulting modified BAC construct contained endogenous promoters linked to hCD2ΔT and rCD2 in the TCRα and Dad1 orientations and is called Δ5’I14 h/rCD2. The Δ5’I14h/rCD2 BAC construct was digested with Not I to liberate the insert from the BAC vector.

2.2 Transgenic Mice

These transgenic animal studies have been reviewed and approved by the Hunter College Institutional Animal Care and Use Committee.

DNA fragments for microinjection were prepared as previously described (Ortiz et al. 2001) except for the BAC construct. Here, the Δ5'I14h/rCD2 BAC was digested with NotI. The preparative digestion products were separated by field inversion gel electrophoresis (FIGE–BioRad), the ~80kb fragment was excised from the gel then purified by dialysis followed by phenol/chloroform extraction.

Purified DNA was microinjected in the pronucleus of (C57Bl/6 x CBA) F1 fertilized mouse eggs. Transgenic founders were identified by Southern blot/PCR analysis on tail DNA. Relative copy number was determined for each line by analysis of at least two Southern blots by PhosphorImager quantification (Molecular Dynamics, Sunnyvale, CA). Lines were directly compared and analyzed for relative copy number on the same Southern blots using the same probe and enzyme digestion. The signal from the endogenous TCRα locus was used as a normalizing control. Mice heterozygous for the transgene was used for analyses.

2.3 RNA extraction

RNA from mouse organs was prepared as previously described using a single step isolation protocol (Chomczynski and Sacchi 1987). Mouse tissues were dissected of fat and washed with phosphate buffered saline to minimize blood contamination then placed in RPMI containing 5% FBS until homogenized. Alternatively, purified T cells and B cells were isolated using a magnetically activated cell separation system (MACS, Miltenyi Biotec Auburn, CA) prior to RNA extraction.

2.4 Northern blot analyses

5ug RNA/sample was run on a 1% agarose gel and transferred onto a non-charged nylon membrane (Genescreen, Perkin Elmer, Waltham, MA) for northern blot analyses using Quickhyb solution (Stratagene, La Jolla, CA). A 428bp BamHI-NcoI genomic fragment of the human β-globin gene coding region was used as a probe to detect human β-globin transgene mRNA. Human CD2 transgene mRNA was detected using a 0.5kb EcoRV-PstI probe from exon II of the human CD2 gene while HLA-B7 mRNA was detected with a 2kb BglII genomic fragment. To normalize for variation, blots were stripped and reprobed with a 500 bp Sau3AI fragment of TCRα constant region (Cα) or 18S(Ambion, Austin, TX). All probes were labeled with [α-32P]dCTP using a random primer labeling kit (Invitrogen, Carlsbad, CA). Transgene signals were normalized and quantified by PhosphorImager (GE Healthcare, Pittsburg, PA) analysis.

2.5 BioInformatic Sequence Analysis

The mouse TCRα LCR region of HS4 (nucleotides 2162 to 3283 of Accession # AF000941) and HS6-316 (Accession # AY545981) was used to identify conserved sequences among vertebrates in the University of California, Santa Cruz (UCSC) Genome Browser, http://genome.ucsc.edu/ (Kent et al. 2002). Sequence for the HS4 and HS6-316 regions were entered in BLAT, the BLAST-Like Alignment Tool (Kent 2002), to find its exact chromosome position, chr 14: 53,188,851- 53,189,980 and chr 14: 53,191,647- 53,191,961 respectively, in the February 2006 Build 36 release of the mouse genome assembly (Waterston et al. 2002). These sequences were then compared with other known vertebrates (human, rat, rabbit) at a base pair level. Conserved sequences were then screened for potential factor binding sites by using TRANSFAC (Matys et al. 2003), http://www.gene-regulation.com/ and ConSite (Sandelin, Wasserman, and Lenhard 2004), http://mordor.cgb.ki.se/cgi-bin/CONSITE/consite.

2.6 Electromobility Shift Assay (EMSA)

Nuclei was prepared as described (Enver, Brewer, and Patient 1985) and resuspended in extraction buffer C described by (Staal et al. 1990). After 30 minutes rocking at 4°C, samples were centrifuged at top speed for 5 minutes. Supernatants were frozen in aliquots at -80°C. 3ug of nuclear extract was used in EMSAs with 50,000 cpm (dry) of 32-P labeled oligonucleotide probe. Samples were incubated on ice 45 minutes with 1.5ug

poly(dI:dC), 1mM DTT, 1mMPMSF, 3% glycerol, 10mM HEPES (pH7.5), 7.5mM NaCl, and 1.25mM KCl. In competition assays, at least 50-fold molar excess of unlabeled oligonucleotide was added. The sequences of oligonucleotide probes used are as follows. HS4-2: ACGTCAGTGATAAGGCAAAATC and ACGTGATTTTGCCTTATCACTG HS6-316: ACGTGGGAAAATCACAACC and ACGTGGTTGTGATTTCCC. Antibodies for supershift assay include mouse anti mouse GATA-3 HG3-31 (Santa Cruz Biotechnology, Santa Cruz, CA) normal mouse IgG (Upstate, Temecula, CA), rabbit IgG (Upstate) rabbit anti mouse p50 (Upstate), p65 (Santa Cruz), or p300CT RW128 (Upstate).

2.7 Flow Cytometry

Single cell suspensions of thymocytes and spleen cells were put in FACS stain buffer (RPMI 1640, 3% FBS, and 10mM HEPES buffer). 10^6 cells in 100uL were blocked, if necessary, for 20 minutes 4°C with an excess of normal mouse IgG. Cells were stained with 0.2-1.0ug of antibody for 20 minutes at 4°C. Three washes in FACS stain buffer was done before flow cytometry analysis. Mouse anti-mouse CD90.2 (Invitrogen) was used to detect splenic T cells. Rat anti-mouse CD19 (BD Pharmingen, San Jose, CA) identified splenic B cells. The appropriate population was gated, analyzed, and compared in both non-transgenic and transgenic cells for hCD2 expression using mouse anti-hCD2 from clone S5.2 (BD Pharmingen). Acquisition and analysis was done using FACScan (Becton Dickinson) and Cell Quest Pro software.

OVERVIEW

Chapter 3: CTCF-independent elements significantly contribute to TCRα locus control region activity

Specific regions of the TCRα LCR must be analyzed to understand their contribution in LCR activity. We analyzed discrete functional elements of the TCRα LCR by making deletion mutants in a transgene-linked LCR mouse model system. Using this model system, we have identified functional regions that gives us a starting point for elucidating the mechanisms of LCR activity.

Chapter 4: Regulatory Influence of the TCRα LCR

The TCRα LCR lies in close proximity to two genes, TCRα and Dad1. TCRα LCR influence on its 5' gene, TCRα, has been studied in vivo. However, recent results in our lab show TCRα LCR linked transgenes lack specific aspects of the TCRα expression patterns that are seen in the endogenous locus. How the TCRα LCR affects the expression of its neighboring 3' gene is unknown. We designed a model system that begins to address how the LCR affects the regulatory pattern of a downstream gene and how the TCRα LCR exhibits its regulatory influence in the context of multiple genes.

Chapter 5: Creating a BAC reporter transgene model to study TCRα/Dad1 regulation

There are other elements besides the TCRα LCR that support expression patterns at the endogenous locus. Because TCRα LCR is insufficient to maintain all the aspects of the endogenous expression profiles, a BAC model system has been designed. This system was made to look at other elements that affect either the 5’ or 3’ genes that flank the LCR and allows us to look at LCR function in a more endogenous-like setting.

CHAPTER 3: CTCF-independent elements significantly contribute to TCRα locus control region activity

3.1 RESULTS

The components of complete TCRα LCR activity are broadly distributed and within a large stretch of DNA sequence. In order to identify important elements that act within it, we looked more closely at several discrete regions. Although none of the regions studied here have classic enhancer activity, previous data indicate that they may play a role in regulating transcription levels and protection against position effect. Analyses of these specific sites were done by making internal deletion mutants and studying mutant LCR activity in a transgenic context. Deletion of these candidate sites in an otherwise functional TCRα LCR reveals information of its role in LCR function in vivo.

3.1.1 Candidate regions for TCRα LCR functional studies

The HS1' sub-element is involved in several ways to ensure proper LCR activity including tissue specific expression. In particular, we and others identified CTCF binding sites shown to participate in enhancer blocking activity within a 94bp region of HS1' (Magdinier, Yusufzai, and Felsenfeld 2004; Gomos-Klein et al. 2007). These are the only CTCF binding sites found in the TCRα LCR.

HS4 is a likely important functional site because of its differential methylation status in lymphoid (thymus and spleen) versus non-lymphoid tissues (Santoso, Ortiz, and Winoto 2000). When a transgene is linked to HS2-6, widespread expression is observed and HS4 is not demethylated in thymus and spleen (Santoso, Ortiz, and Winoto 2000). Previous studies show that the normal methylation patterns are established early in development before recombination occurs. However, HS4's role and function in LCR activity remains unknown.

Another discrete sub-element within HS6 was found to be necessary to suppress position effect. When HS6 was used in isolation, linked transgenes can overcome position effect with widespread activity. It can even overcome position effects in non-lymphoid cells, such as NIH3T3 cells. In these cell culture studies, a deletion of just 316bp of the 1.8kb HS6 region caused an increased incidence of position effect (Harrow et al. 2004). A logical next step for this finding was to test this 316bp region in the context of the entire TCRα LCR in a transgenic model.

3.1.2 Transgenic mouse models

Much of our understanding of the components of TCRα LCR function has come from analyses of reporter transgene loci in mice. This study used a 4.9-kb human β-globin reporter gene fragment linked to either the wild type TCRα LCR (named β:1-6) or

internal deletion mutants thereof. The β-globin reporter fragment contains the human β-globin promoter, exons, introns and 3' enhancer sequence. It is highly subject to position effect silencing in the absence of additional elements (Magram, Chada, and Costantini 1985; Townes et al. 1985) and has long been used as a reporter of LCR activity in multiple systems (Ortiz et al. 1997; Chauveau et al. 1999; Grosveld et al. 1987; Greaves et al. 1989). Including an intact LCR in the transgene eliminates these position effects. Linking the β-globin reporter gene to HS1-6 of the TCRα LCR yields the high-level, lymphoid organ-specific and relatively consistent reporter expression levels per copy (within a narrow 2-3 fold range) indicative of integration-site independence and full TCRα LCR activity in vivo (Ortiz, Cado, and Winoto 1999).

Deletion mutant transgenic lines were designed and compared to transgenic mice linked to a fully active HS1-6 TCRα LCR. Deletions in the context of TCRα LCR HS1-6 removed either the CTCF binding sites, HS4, or HS6-316 regions described above (Figure 5). DNA fragments containing these constructs were purified for microinjection into mice. The resulting transgenic mice provided our model system with which to analyze LCR function.

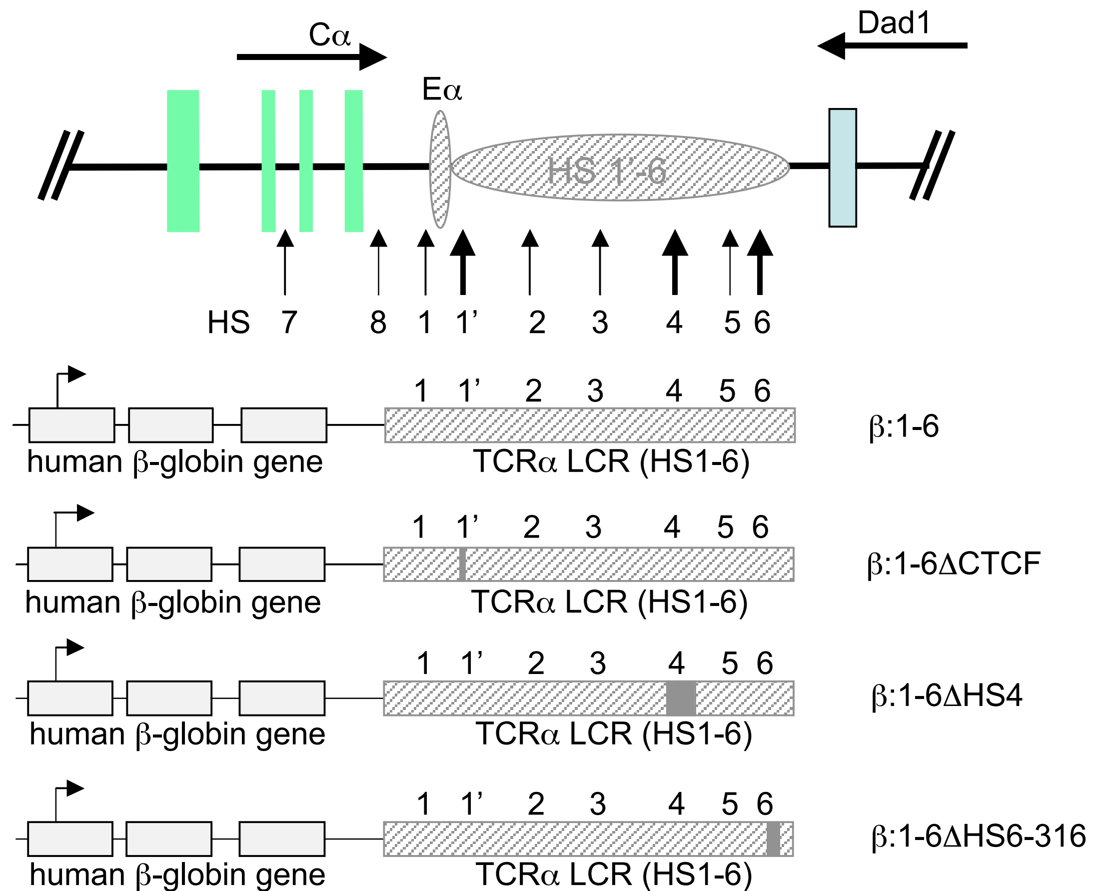


Figure 5. TCRα LCR and transgene constructs. Top: diagram (not drawn to scale) of the 3' end of the genomic locus. Vertical arrows and numbers indicate the 9 DNase1 HS of the LCR. The hashed area indicated the minimal required DNA sequences for full LCR activity. Horizontal arrows indicate the transcription orientation of genes. Transgene constructs: linked the human β-globin reporter gene 5' of the indicated HS of the LCR. The filled boxes indicate the region within the LCR that is deleted for each mutant transgene. Figure reprinted from Gomos-Klein et al. *Copyright 2007. The American Association of Immunologists, Inc.*

Twenty independent transgenic lines containing one of these four constructs were

analyzed. For each of these lines, copy number was determined by multiple Southern blot analyses normalized to the endogenous two copies of TCRα. Six independent transgene lines containing β-globin linked to HS1-6 were analyzed and called β:1-6 lines 14, 32, 34, 35, 42 and 58 with copy numbers of 16, 4, 7, 9, 10 and 3 respectively. Four separate β:1-6ΔCTCF lines 2,3,11 and 16 were created for the deletion mutation in CTCF. Their copy numbers were 12, 4, 10 and 11. There were four independent lines of HS4 deleted transgenic mice β:1-6ΔHS4, Lines 26, 40, 50 and 64 with 1, 1, 5 and 14 transgene copies respectively. Finally, the deletion in HS6-316 yielded six separate β:1-6ΔHS6-316 lines 5, 20, 26, 49, 50 and 60 with copy numbers of 9, 38, 5, 6, 17 and 6.

3.1.3 Transgene expression levels due to LCR deletion mutations

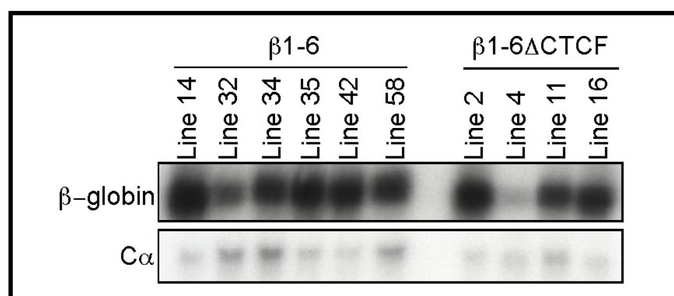
LCRs normally confer copy number-related expression upon a linked transgene. So, we examined whether transgene expression levels were altered when driven by the various deletion mutants.

RNA was extracted from thymic tissues. Expression levels of each of the transgenic lines were normalized to the endogenous TCRα in thymus. Normalized expression levels were then divided by copy number for each transgenic line. Expression levels per copy were quantified by northern blot and PhosphorImager analyses.

CTCF sites in the TCRα LCR contributes nominally to expression levels:

Expression levels per copy in the thymus of the fully active LCR-linked transgenics, β:1-6, and the CTCF deletion mutants, β:1-6ΔCTCF, were compared. A representative Northern blot is shown in Figure 6A and quantified expression levels per copy are shown in Figure 6B. The data from this representative experiment show the average transgene expression levels per copy from the ΔCTCF mutant construct are slightly reduced from the average of wild type β:1-6 transgene expression levels. Average β:1-6ΔCTCF expression is 77% of wild type expression in thymus. Thus, it appears that the tandem CTCF sites are minor contributors in the TCRα LCR for expression of a linked transgene.

A.



B.

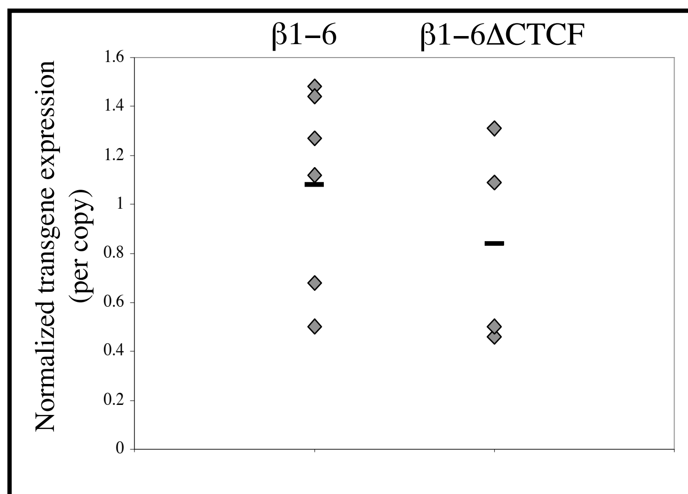
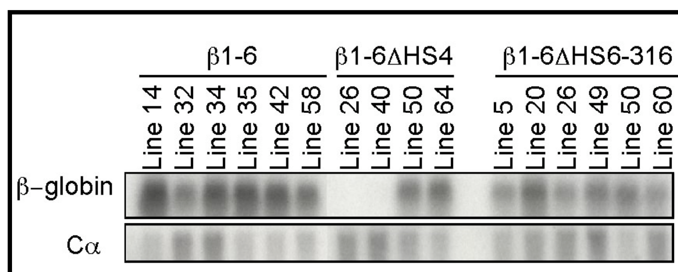


Figure 6. Minor effect of transgene expression levels per copy due to deletion of CTCF sites in the TCRa LCR. A, Representative northern blot analyses of reporter transgene expression in thymus of multiple lines of mice bearing either the wild-type (b:1-6) or CTCF site region deleted (b:1-6ΔCTCF) LCR-driven reporter constructs. The northern blot shows β-globin reporter gene expression and the coordinating endogenous TCRα expression. The relative copy numbers from left to right are: 16, 4, 7, 9, 10, 3, 12, 4, 10, 11. B, PhosphorImager analyses of northern blot plotted as normalized transgene expression per copy. TCRa (Ca) mRNA was used as a normalizing control for thymus samples. The gray diamonds indicate the expression level per copy in each individual line. The horizontal dash indicates the average of the expression levels per copy of all transgenic lines bearing the same construct, as indicated. Figure reprinted from Gomos-Klein et al. *Copyright 2007. The American Association of Immunologists, Inc.*

HS4 and HS6-316 deletions in the TCRα LCR impairs thymic expression:

Using the same methodology described above, we quantified the contribution of the HS4 and HS6-316 elements to TCRα LCR driven transgene expression in thymus. Average expression in the thymus for these mutants are much lower compared to the WT control. We compared transgene mRNA levels (per copy) produced by the β:1-6ΔHS4 and β:1-6Δ316 transgenes to those of the wild type LCR driven β:1-6 construct. In the representative experiment (Figure 7), average transgene mRNA levels per copy are reduced over 4-fold from wild type levels when HS4 DNA is absent from the LCR. Similarly, average expression levels of the β:1-6Δ316 mutants are 3.3-fold lower per transgene copy, compared to that driven by the wild type LCR. These data help confirm that the HS4 and HS6-316 regions of the TCRα LCR harbor in vivo functional elements.

A.



B.

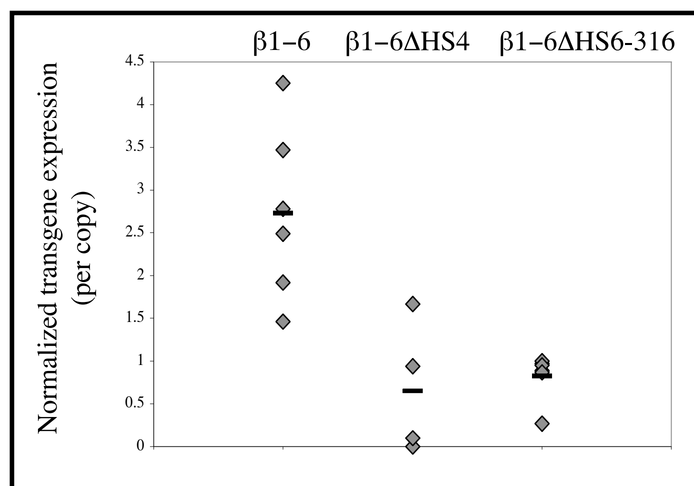


Figure 7. Transgene expression levels per copy are significantly reduced upon internal deletion of either the HS4 or HS6-316 regions. A, Representative Northern blot analyses of reporter transgene expression in thymus of multiple lines of mice bearing either the wild-type (β:1-6), HS4 deleted (β:1-6ΔHS4), or HS6-316-deleted (β:1-6ΔHS6-316) LCR-driven reporter constructs. B, PhosphorImager analyses of Northern blots plotted as normalized transgene expression per copy. TCRα (Cα) mRNA was used as a normalizing control for thymus samples. The gray diamonds indicate the expression level per copy of the indicated organ in each individual line. The horizontal dash indicates the average of the expression levels per copy of all transgenic lines bearing the same construct, as indicated. Figure reprinted from Gomos-Klein et al. *Copyright 2007. The American Association of Immunologists, Inc.*

Functional regions of the TCRα LCR that affect expression:

Using this data and others not shown here, we compared lymphoid organ (both thymus and spleen) expression levels of the deletion mutants to WT expression. Figure 8 is a compilation of what is known about lymphoid specific expression due to the TCRα LCR. Although each of the regions that have been studied affect LCR activity, we found that a deletion in HS4 or HS6 decreased expression levels compared to WT by 60-70%. The regions tested are necessary to drive LCR expression in lymphoid organs even though they have no classic enhancer activity.

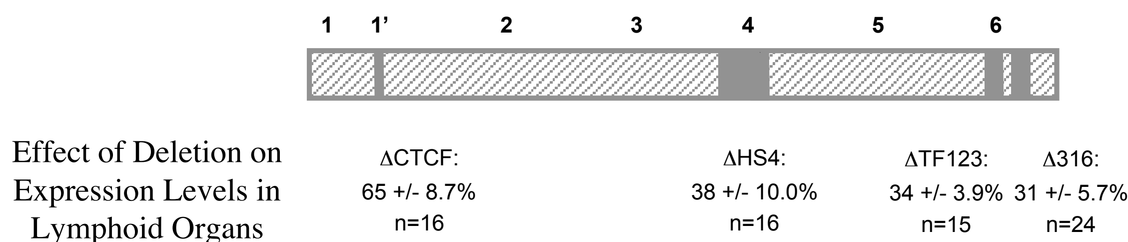


Figure 8. Summary of data from multiple experiments showing the effect of internal deletion of the indicated DNA sequences on TCRα-driven expression levels (per copy) in lymphoid organs (thymus and spleen) downstream of HS1. Filled boxes represent deleted sequences. The average percentage of wild-type LCR-driven expression (+/-SE) observed in the indicated deletion mutant is shown. ΔTF123 was previously described (Harrow et al. 2004; Ortiz et al. 2001) and is included in this summary for a more comprehensive presentation of active regions downstream of the Eα/HS1 region. n represents the number of mutant LCR-driven data points used to calculate the average. Figure reprinted from Gomos-Klein et al. *Copyright 2007. The American Association of Immunologists, Inc.*

3.1.4 Elements that contribute to position effect suppression

Normally, linking a reporter gene to the TCRα LCR yields lymphoid organ specific and consistent reporter gene expression patterns. Relative to the thymus, splenic reporter gene expression is between 20-40% while non-lymphoid organ expression is less than 10%. The TCRα LCR directs this pattern of expression very consistently regardless of the transgene's random integration site (Harrow et al. 2004; Ortiz, Cado, and Winoto 1999; Diaz, Cado, and Winoto 1994).

mRNA expression levels of the reporter transgenes in various lymphoid and non-lymphoid tissues were analyzed. Specifically, we measured expression levels of thymus and spleen lymphoid tissues. Non-lymphoid tissues included in this study were kidney, lung, liver, and heart. Each of the four to six independent lines per construct were normalized to 18S rRNA then graphed relative to thymic expression. Thymus is where reporter transgene expression should be the highest. Line-to-line consistency in relative tissue distribution for each construct was compared to the β:1-6 transgenic lines. Significant deviation from this consistent pattern indicates regulating interference at the site of integration. In these cases, transgene expression is subjected to varying native regulatory sequences, i.e. position effect, which indicates impaired LCR activity.

Consistent with previous data, the six new independent β:1-6 transgenic lines containing the fully active LCR consistently displayed the expected expression pattern. The expression profile relative to thymic expression displayed very little variation between

independent lines (Figure 9A). This consistency is most evident by the small standard deviation in each of the tissues analyzed. As expected, non-lymphoid organ transgene expression was maintained at less than 10% relative to thymus.

The effect of deleting the CTCF sites from the LCR on the normal transgene expression pattern was minor with mRNA levels only barely crossing the 10% “threshold” (relative to thymic mRNA levels) in lung and heart (Figure 9B).

However, we found that HS4 and HS6-316 elements are necessary to protect its linked transgene from position effects. Tissue distribution of transgene expression was more perturbed in the absence of the 316-bp region of HS6 with average relative expression levels in some non-lymphoid organs exceeding 20% of that seen in thymus (Figure 9C). Line to line consistency within this β:1-6ΔHS6-316 was reduced as evidenced by the larger standard deviation in each of the tissues relative to β:1-6 expression. Finally, the relative tissue-distribution of multiple lines bearing the β:1-6ΔHS4 mutant transgene was plainly abnormal leading to large variations in organ expression levels relative to thymus among the four lines analyzed (Figure 9D). There was little to no consistency of expression between lines of this construct.

Inconsistent tissue distribution is a manifestation of integration-site dependent position effects. These data indicate that the mutant LCR in β:1-6ΔHS4 and β:1-6ΔHS6-316 were unable to reliably establish its normal expression program at some integration sites. Our interpretation is that β:1-6ΔHS6-316 and β:1-6ΔHS4 mutant LCRs are less able to protect

the transgene from position effects. Ability to overcome regulatory influences at a random integration site is a characteristic of insulators. Consistent with our results, prior data from other labs suggested that the TCRα LCR contains multiple insulator-like regions (Zhong and Krangel 1999; Magdinier, Yusufzai, and Felsenfeld 2004) whose precise location and contribution to LCR activity is still not known.

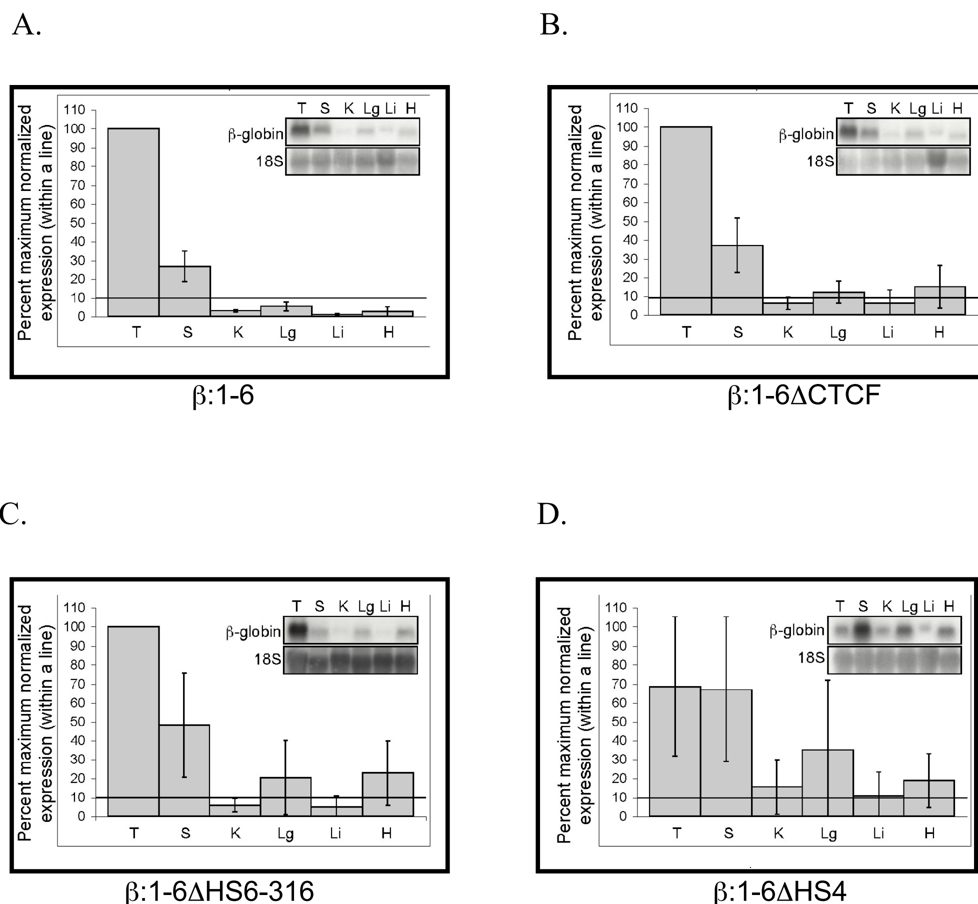


Figure 9. Line-to-line consistency in relative tissue distribution of transgene expression from wild-type and mutant LCR-driven constructs. A, PhosphorImager analyses of Northern blots of organ RNA from six independent transgenic lines bearing the β:1-6 (wild-type LCR-driven) reporter construct The inset is a representative Northern blot on β:1-6 line 42 showing the reporter β-globin signal and 18S rRNA signal used as a normalizing control. The graph represents the average relative tissue distribution of normalized transgene expression in the indicated organs (T, thymus; S, spleen; K, kidney; Lg, lung; Li, liver; H, heart) among the various lines. Within each line examined, thymic transgene expression was designated as 100%. Reporter expression from other organs within the same line was plotted as a percentage of thymic expression. Error bars, The SD in the percent thymic expression levels for each organ among the six β:1-6-transgenic lines. Note the low variation in percentage of thymic expression from line-to-line and that nonlymphoid expression is well below the 10% of maximum level marked by the horizontal bar. B, Similar analyses of four independent lines of transgenic mice bearing the β:1-6dCTCF construct with the Northern blot from β:1-6ΔCTCF line 16 shown in the inset. C, Similar analyses of six independent transgenic lines bearing the β:1-6ΔHS6-316 construct with the Northern blot from β:1-6ΔHS6-316 line 20 shown in the inset. D, Similar analyses of four independent transgenic lines bearing the β:1-6ΔHS4 construct with the Northern blot from β:1-6ΔHS4 line 40 shown in the inset. Note the greater variation in relative tissue distribution of transgene expression seen in C and D as compared with A and B. Figure reprinted from Gomos-Klein et al. *Copyright 2007. The American Association of Immunologists, Inc.*

3.2 SIGNIFICANCE

3.2.1 Discrete functional sub-regions of the TCRα LCR identified

Lymphoid organ (both thymus and spleen) expression levels of the deletion mutants indicate that each of the regions that have been studied affect LCR activity at varying levels. None of the regions tested contain classic enhancer activity. Yet, they are necessary for full LCR activity in lymphoid organs. The CTCF sites in HS1’ seems largely dispensable for maintaining lymphoid expression levels. This supports that its regulatory function may exist in supporting other aspects of LCR activity as HS1’ exists in both lymphoid and non-lymphoid tissues. HS4 and HS6 are necessary for high levels of expression. These regions might synergize to create a strong positive regulatory influence in lymphoid specific tissues. Alternately, our data seem to indicate that each region, especially at the 3’ end of the LCR, contributes to the strength of LCR insulation capacity.

3.2.2 Insulator-like activity of HS4 and HS6-316

Unchecked spreading of condensed chromatin is one likely cause of the position effect gene silencing often observed from randomly integrated transgenes in mice (Palmiter and Brinster 1986; Festenstein et al. 1996; Milot et al. 1996). LCR activity enables an

ectopically integrated transcription unit to establish its own expression program independent of the regulatory influences of its site of integration. As such, LCR activity resembles the poorly understood genomic insulation mechanisms that establish discrete regulatory domains.

The study of the components of LCR position effect suppression activity may yield insight into the broader question of how differential gene regulation programs can be organized in complex genomic loci. The TCRα LCR is native to a genomic location flanked by differently regulated genes (TCRα and Dad1). It seems likely that there exist a variety of cis-acting elements that can serve a barrier function separating the regulation of neighboring genes. Therefore, its DNA is likely to contain important examples of elements playing a role in the processes limiting cis-regulatory influence in the genome.

Insulators are a type of cis-regulatory element proposed to help establish autonomy of gene loci. In functional assays, insulators display enhancer-blocking activity and/or a barrier capacity with the ability to halt the propagation of repressive chromatin states (Figure 10) [reviewed in (Gaszner and Felsenfeld 2006)]. Thus far, the only functional sequence element common to multiple vertebrate enhancer-blocking-type insulators is the CTCF binding site (Bell, West, and Felsenfeld 1999). However, there exist less well characterized insulators (of both the enhancer blocking and barrier types) whose activities are clearly CTCF-independent (Magdinier, Yusufzai, and Felsenfeld 2004), reviewed in (West, Gaszner, and Felsenfeld 2002). The TCRα LCR we study has both CTCF-dependent and CTCF-independent enhancer- blocking activities (Magdinier, Yusufzai,

and Felsenfeld 2004; Zhong and Krangel 1999). Therefore, our experimental model provided us an opportunity to further characterize these various insulator elements and their roles in gene regulation in vivo, using an integration-site independent, LCR-driven reporter transgene system. Seemingly conflicting data have been reported regarding the role of CTCF and its cognate binding site in position-effect suppression in chromatin. Work on the chicken β-globin insulator has produced clear evidence that its barrier capacity is independent of its CTCF binding sequences (Recillas-Targa et al. 2002; West et al. 2004). However, reports using other experimental systems have asserted that CTCF can block the spreading of repressive chromatin (Defossez and Gilson 2002), regulate the balance between activating and silencing histone modifications (Splinter et al. 2006) and protect from position effects (Filippova et al. 2005).

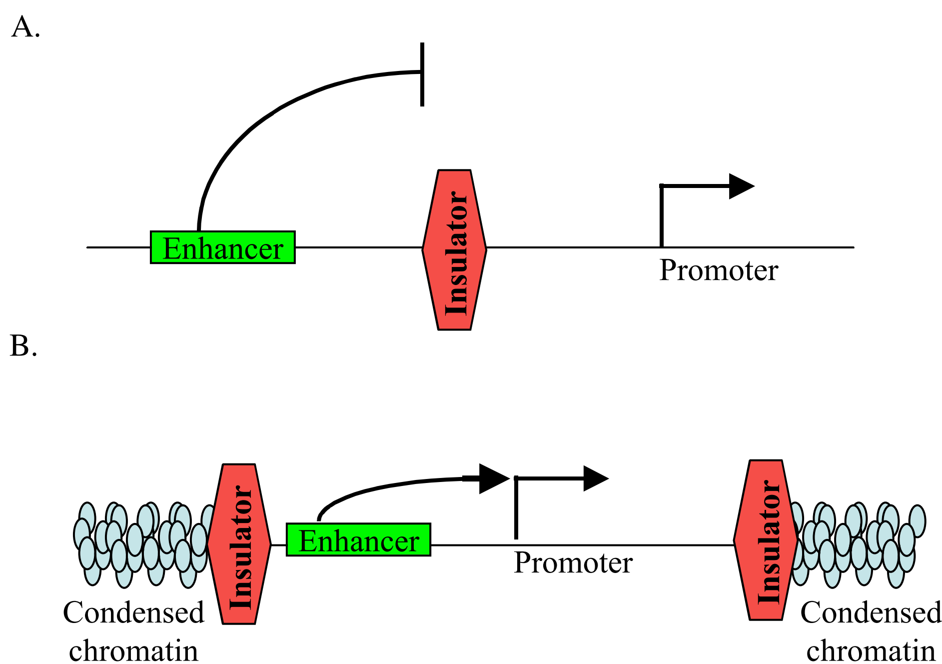


Figure 10. Insulators can help define a regulatory region. A. They can block enhancer-promoter interaction (enhancer blocking), and/or B. Insulators can prevent the spread of condensed chromatin (barrier).

We found that HS4 and HS6 are required for the LCR to support complete position effect suppression while a deletion of the tandem CTCF-like binding sites did not have significant impact on TCRα LCR position suppressing activity in vivo. HS4 and HS6 do not have CTCF binding sites. Previous studies have shown that broad regions 3' of the CTCF sites are involved in insulator activity (Zhong and Krangel 1999; Magdinier, Yusufzai, and Felsenfeld 2004). But, this is the first report identifying discrete elements within this putative insulating area as regions needed to protect against position effect in a transgenic context.

HS4 and HS6-316 may well be two examples of CTCF-independent barrier-type insulator elements. These elements share some characteristics with known barrier insulators. For example, in T cells, the endogenous HS4 region DNA shows a higher level of histone acetylation than its neighboring sequences (Magdinier, Yusufzai, and Felsenfeld 2004), a trait of the chicken β-globin insulator (Litt et al. 2001; Mutskov et al. 2002). Also, the HS6- 316 region protects stable-transfected reporter constructs from silencing in chromatin, the hallmark property of barrier insulators (Felsenfeld et al. 2004). It has been proposed that the non-enhancer components of LCRs might have barrier insulator properties (Gaszner and Felsenfeld 2006). There are multiple position effect suppressing elements in the TCRα LCR, providing evidence to support a proposed model of barrier insulator activity (Gaszner and Felsenfeld 2006). In this model, barrier function depends on the local balance between negative and positive regulatory activity in chromatin. The functional elements at the 3' end of the TCRα LCR likely synergize to create a strong barrier to prevent position effect silencing. We have shown that removal

of either one of these elements reduces the strength of the barrier allowing a greater incidence of position effects.

These data indicate that the enhancer blocking activity demonstrated for the CTCF-binding region does not meaningfully contribute to overcoming heterochromatin induced position effects in our system. Overcoming these effects is a hallmark of LCR activity at ectopically integrated transgene loci (Festenstein et al. 1996; Milot et al. 1996). However, the enhancer blocking data reported previously (Magdinier, Yusufzai, and Felsenfeld 2004; Zhong and Krangel 1999) suggest that these sites are not entirely non-functional. It is possible that the CTCF sites act as an enhancer blocker that is only active in cases where the activity of the TCRα gene must be protected from the regulatory influence of the Dad1 gene nearby, or vice versa. TCRα and Dad1 genes maintain very different expression profiles during T cell development and in various tissues (Hong et al. 1997; Hong et al. 1999).

Although there is functional resemblance of position-effect suppression activity in LCRs and barrier insulator activity, it is apparent from the literature that several LCR elements contribute to its complex activity including for barrier function. Our data to date on the TCRα LCR indicate that this complexity may arise from the existence of multiple position effect suppressing elements in the 3' region of the LCR and their interactions with the critical LCR elements previously identified in its more TCRα - proximal 5' portion (Ortiz, Cado, and Winoto 1999). Having now localized these position effect suppressing elements, future studies can be aimed at the important question of what

mechanism drives the LCR to collectively establish the independence of a gene locus.

3.3 FUTURE DIRECTION

3.3.1 Conserved Sequences in HS4 and HS6-316

An obstacle to identifying the molecular basis for the position effect suppressing activity of LCRs has been the puzzling lack of sequence homology between LCRs, even among their non-enhancer functional regions identified to date. To gain insight into the sub-sequences likely to be important in the function of the position effect suppressing elements identified in this study, we examined cross-species DNA sequence homology in the HS4 and HS6-316 regions using the UCSC genome browser (Figure 11). We found several sub- sequences within these elements that are highly conserved among species. Many of these conserved sequences contain consensus recognition sites for known transcription factors.

HS4:

Mouse	AATTAGAATGGCAAACCCGAGC
Rat	GATTAGAATGACAAAACAGAGC
Human	GATCAGAATGGCAAGGCAGAGC
	c-Rel family
Mouse	GTCAGAAGTTCAGTGATAAGGCAAAATCAGAAGA
Rat	GTCAAAAGTTCAGTGATAAGGCCAAATCAGAAGA
	GATA-3

HS6-316:

Mouse	CAAATTTGAACTCA	Mouse	GGGAAAATCACAACC
Human	TAAATCTAATCCCA	Human	GGGTAAATCACAACC
Rabbit	TAAATTTAAATCCT	Rabbit	GTGTAAATCACAACC
			c-Rel family

Figure 11. Conserved DNA sequences exist within HS4 and HS6-316. Sequences of mouse HS4 and HS6-316 were compared and aligned with other mammals (indicated at left) using the University of California Santa Cruz (UCSC) Genome Browser (see Materials and Methods). Consensus recognition sequences for transcription factors of the c-Rel family and GATA-3 are underlined.

For example, the HS4 element contains a 38-bp sequence with 92% identity between mouse and rat loci. This sequence features a binding site for the GATA-3 transcription factor. GATA-3 is one of the most critical factors in T cell development as GATA-3 null mice have a specific defect in T- lineage development (Ting et al. 1996). GATA-3 plays a role in TCR gene function with high levels of expression at stages of T cell development where the TCRα LCR is highly active (Hernandez-Hoyos et al. 2003).

GATA-3 also induces chromatin remodeling at the TH2 cytokine locus (Takemoto, Arai, and Miyatake 2002) and Interleukin-10 locus (Shoemaker, Saraiva, and O'Garra 2006) during T helper cell differentiation.

HS4 DNA also contains a second conserved 22-nucleotide stretch with 73% identity between mouse, rat and human loci that bears a site for binding of rel family proteins. In the HS6-316 element, a 15-bp sequence 86% conserved among mouse, human and rabbit also contains a rel-consensus recognition site. The rel family of transcription factors is involved in lymphocyte responses, proliferation and survival. A rel-family dependent cis-element was found to be required for chromatin modification at the Macrophage Chemotactic Protein (MCP)-1 locus (Teferedegne et al. 2006). These are properties similar to that defined for the barrier elements of the chicken β-globin insulator where specific factor binding sites were found to attract histone modification enzymes to a gene locus to protect against position effect silencing (West et al. 2004).

3.3.2 Binding sites in HS4 and HS6-316

Electromobility Shift Assays (EMSA) can detect the interaction of protein factor complexes with DNA. EMSA competition experiments can be a tool to preliminarily identify candidate transcription factor families that the complex may bind. In this assay, active transcription factors bind to their consensus sequence or to the highly conserved regions identified above.

Using the bioinformatics results above, I created oligonucleotides either containing the highly conserved regions or consensus sequences to the factor binding sites that were predicted. Oligonucleotides were incubated with thymic nuclear extract. Preliminary data show that identified conserved regions in HS4 and HS6-316 have DNA binding sites that form a complex (Figure 12). Specificity of binding sites in a conserved HS4 region containing the putative GATA-3 site was observed by competition assays against molar excess of unlabelled competitors, including itself and an 18-bp GATA-3 consensus sequence (data not shown). A supershift using an antibody to GATA-3 in the conserved HS4 is shown in Figure 12A. Similarly, specificity of binding sites in a conserved HS6-316 region was shown by competition assays against itself, p50, and IgkB (data not shown). A partial supershift indicating p50 interaction in HS6-316 was observed (Figure 12B).

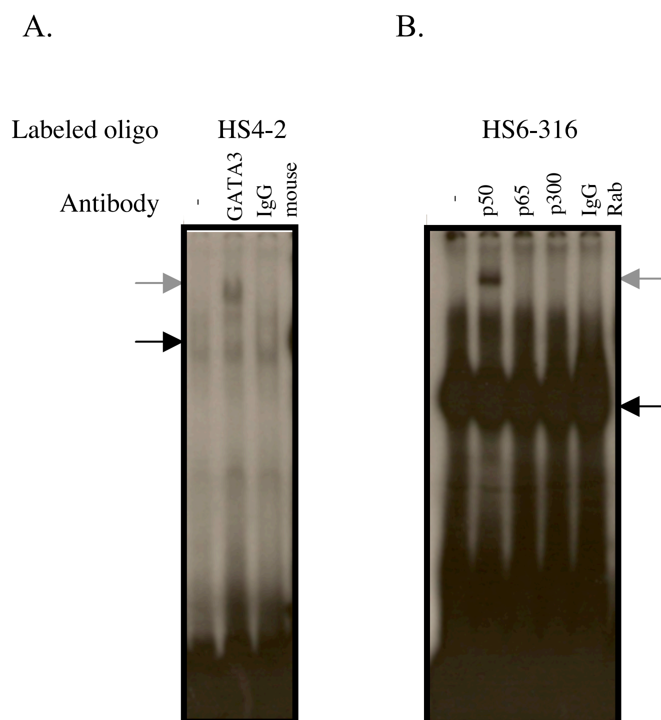


Figure 12. DNA binding sites within conserved regions of HS4 and HS6-316. Preliminary EMSA experiments showing formation of complexes with conserved HS4 and HS6-316 regions. Black arrows indicate complex formation with the conserved sequence. Grey arrows indicate supershift with specific antibodies indicated. A. Conserved sequence of HS4-2 predicted binding site indicates complex formation with GATA-3. B. Strong complex formation at conserved and predicted binding site is indicated by black arrow. Among factors that interact with this complex is p50.

3.3.3 In vivo footprint

Other methods can be used to detect the important functional sequences in these regions. In order to definitively detect factor binding sites, in vivo footprint can be done of the HS4 and HS6 regions. In vivo footprinting identifies factor occupancy of genomic DNA. The DNase I in vivo footprinting assay is a method of studying DNA-protein interaction

and identifies the DNA sequence to which a protein binds in specific tissues.

3.3.4 Isolate sequence conferring insulation capacity

HS6 was found to display an insulator-like function in stable-transfected murine NIH3T3 cells (Harrow et al. 2004). HS6-316 was shown to be an important functional component of the protection from position effect in these studies (Harrow et al. 2004). The next step in these studies is to use established cell culture enhancer blocking and barrier insulation assays to confirm and further dissect the insulation capacity of specific sequences in HS4 and HS6 (Zhong and Krangel 1997, 1999; Recillas-Targa et al. 2002).

CHAPTER 4: Regulatory Influence of the TCRα LCR

4.1 RESULTS

TCRα LCR influence on its 5' gene, TCRα, has been studied in vivo. However, recent results in our lab have shown TCRα LCR linked transgenes lack aspects of all the TCRα expression patterns seen in the endogenous locus (Harrow and Ortiz 2005). This may be due to other unidentified elements in the locus or a level of regulation that cannot be emulated in a simple, single gene environment. At its native locus, the TCRα LCR lies in close proximity to two genes, TCRα and Dad1. How the presence of two flanking genes in close proximity influences the TCRα LCR's activity has not been studied. In addition, TCRα LCR influence on the expression of its neighboring 3' gene is unknown.

In order to look at the regulatory control of TCRα LCR in specific tissues and cell types, the lab created a new transgenic model system using human CD2 (hCD2) and human leukocyte antigen-B7 (HLA-B7) as reporter genes. Reporter gene constructs contained one or two reporter genes flanking TCRα LCR. This 'double reporter transgene' model more closely resembles the endogenous locus. Given that the endogenous LCR lies between differentially expressed genes, we wanted to access the regulatory significance of gene position relative to the LCR. This model system begins to address how the TCRα LCR affects the regulatory pattern of a downstream gene and what regulatory role it may play in the context of multiple genes.

4.1.1 Transgenic mouse models

hCD2:1-8 (hCD2)

Previously reported by our lab is a transgenic model that allowed for single-cell analysis of TCRα LCR activity (Harrow and Ortiz 2005). The construct for these transgenic lines used an hCD2ΔT reporter gene (Melton et al. 1996) linked to the wild type TCRα LCR (Figure 13). This reporter gene has an early stop codon inserted that allows for cell surface expression but prevents normal intracellular signaling associated with hCD2 (Melton et al. 1996). Without its LCR, hCD2ΔT is poorly expressed and subject to integration position effects (Lang et al. 1991; Festenstein et al. 1996). Three independent mouse lines containing this construct called hCD2 were used for this study: lines 4, 29, and 44. Copy number was determined to be 10, 5, and 5.

1-8:HLA-B7 (B7)

The HLA-B7 gene is ubiquitously expressed as a human class I major histocompatability (MHC) α chain and is involved in antigen presentation. Like hCD2, expression of HLA-B7 requires additional regulatory elements for proper expression (Kushida et al. 1997). The 6.5kb HLA-B7 fragment previously described (Chamberlain et al. 1988) was linked to the 3' portion of the TCRα LCR. To further emulate the endogenous situation, HLA-B7 was also placed in the transcription orientation of Dad1 (Figure 13). This construct, called B7, yielded 3 independent transgenic lines (13, 40, and 50) with copy numbers of 3, 3, and 7.

CD2:1-8:B7 (CD2:B7)

To further analyze the regulatory role of the LCR on its neighboring genes, a transgene construct was made that more closely emulates the endogenous locus with 2 reporter genes, hCD2 and HLA-B7, flanking the LCR (Figure 13). The hCD2 reporter gene described above was placed 5' of the TCRα LCR while the HLA-B7 transgene was put in the Dad1 orientation 3' of the TCRα LCR. These double reporter transgene lines were called CD2:B7. Three independent lines were established (lines 10,17,22) with copy numbers of 5,5, and 10.

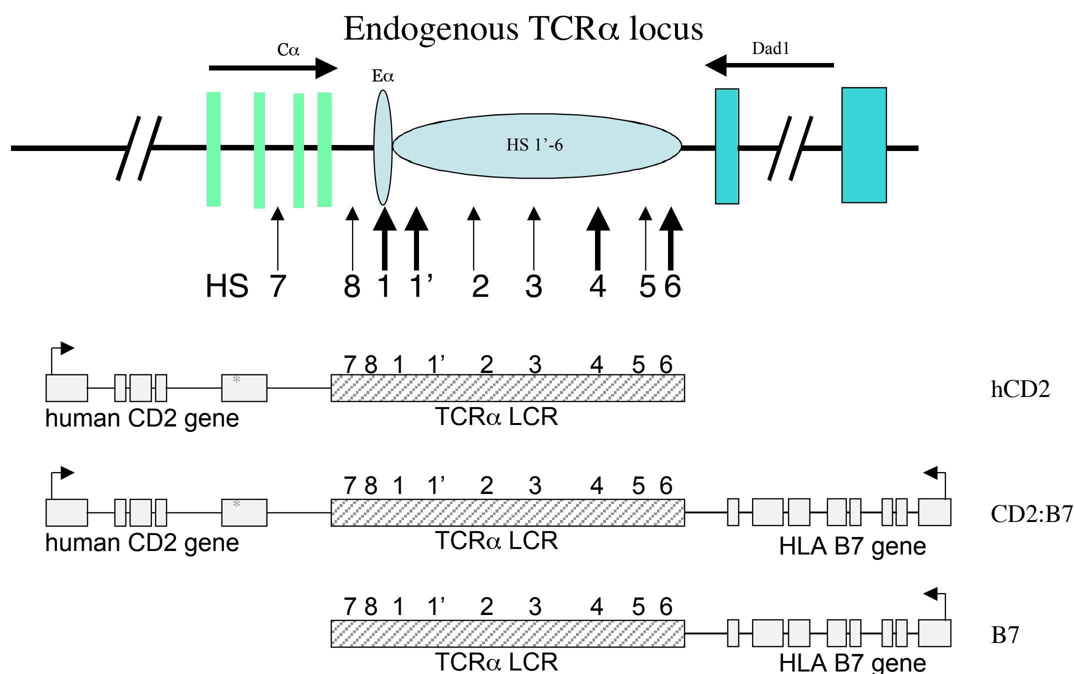


Figure 13. TCRα LCR and transgene constructs. Top: diagram (not drawn to scale) of the 3' end of the genomic locus. Vertical arrows and numbers indicate the 9 DNase1 HS of the LCR. Horizontal arrows indicate the transcription orientation of genes. Solid boxes indicate exons. '*'indicates early stop codon. Transgene constructs: linked human CD2 reporter gene 5' of the indicated HS of the LCR, HLA-B7 linked 3' of the LCR or a combination of both reporter genes (construction of transgenes by Faith Harrow).

4.1.2 Transgene expression levels are related to copy number regardless of TCRα LCR orientation

One of the hallmarks of LCR activity is a linear correlation of expression with copy number in transgenic mice. Reporter gene expression was measured using whole thymic RNA by northern blot. Expression was measured using probes for either hCD2 or HLA-B7. Expression was then normalized to endogenous TCRα expression levels. This value was then divided by copy number. The single hCD2 and B7 constructs were measured for their reporter gene expression while CD2:B7 transgenic lines were measured for both hCD2 and HLA-B7 reporter genes. We compared hCD2 and HLAB7 expression from the dual reporter transgene system to hCD2 and B7 single-gene reporter systems in multiple lines of transgenic mice.

Interestingly, expression was still related to copy number regardless of orientation to the LCR. hCD2 expression in various lines were within 3.2-fold of each other while HLA-B7 expression per copy fell within a narrow 2.3- fold range (Figure 14).

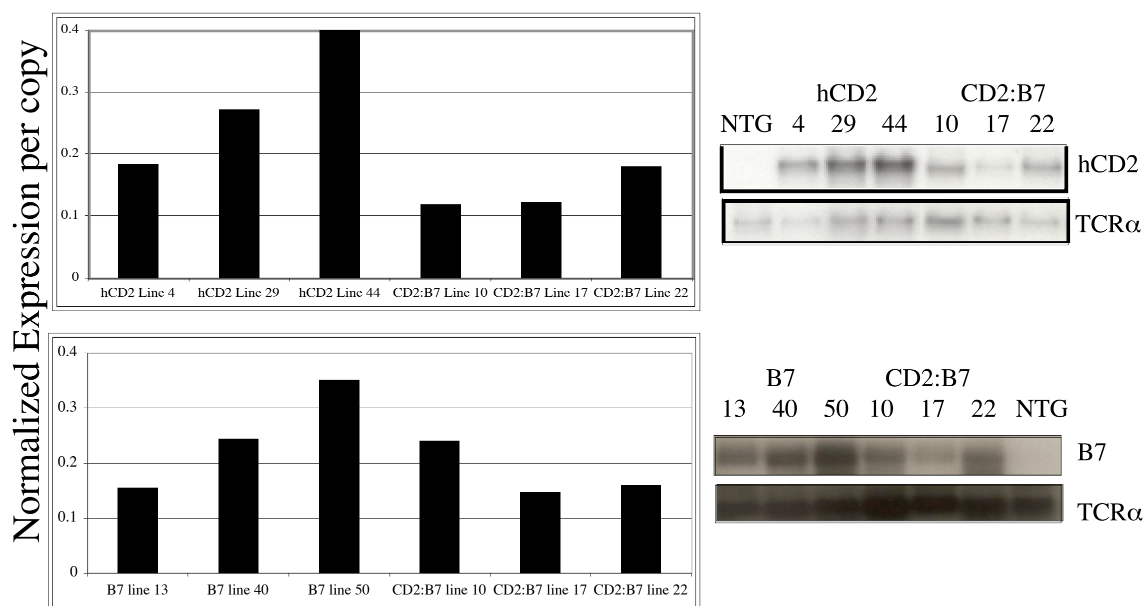


Figure 14. Transgene expression correlates to copy number regardless of orientation. Representative northern blot and quantification of expression of hCD2 and B7 mRNA in the thymus on a per copy basis. Expression was measured by northern blot and quantified using endogenous TCRα as a normalizing control. All the reporter genes placed 5', 3', or flanking the LCR exhibited expression in a linear, copy number dependent fashion within 3.2-fold or less.

4.1.3 Transgenes flanking either side of the TCRα LCR display TCRα- like expression

TCRα is normally expressed in lymphoid tissues while Dad1 is ubiquitously expressed. Expression levels of both reporter transgenes were investigated in a variety of tissues. RNA was extracted from whole organs of lymphoid tissue (thymus and spleen), as well as non- lymphoid tissues (kidney, lung, liver, and heart). In order to compare expression patterns among each of the various constructs, the normalized percent expression of each

tissue relative to thymus within each individual mouse was calculated. Three independent lines bearing each construct was analyzed. The results from these lines were averaged in order to compare line-to-line consistency and to be able to analyze gross effects on tissue distribution of transgene expression due to reporter gene position with respect to the LCR.

As expected, hCD2 followed a consistent TCRα expression profile for both constructs containing hCD2 in the TCRα orientation. The tissue distribution of hCD2 expression in six independent transgenic lines, three from the hCD2 construct and three from the CD2:B7 construct, were analyzed relative to thymus (Figure 15A,B). Interestingly, we found that the tissue distribution expression pattern for HLA-B7, whose location and orientation mimic the Dad1 gene, also followed a TCRα-like expression pattern (Figure15B,C). Six independent lines from either CD2:B7 or B7 constructs, showed consistent TCRα -like expression. It was noted that non-lymphoid expression of B7 is slightly higher than hCD2. The increased levels of expression may be due to the reporter gene promoter. MHC I molecules are present in almost every nucleated cell, thus HLA-B7 can be expressed on all cell types. CD2, on the other hand, is expressed on B cells, T cells and NK cells. Another possibility for the difference of expression may be its position relative to the LCR. This difference in expression levels can be tested by switching the reporter gene orientation whereby HLA-B7 would be placed 5' of the LCR while hCD2 would be placed 3' of it.

It was also noted that relative to thymus, hCD2 expression in the Spleen was decreased in

the CD2:B7 lines compared to the hCD2 lines (Figure 15B). However, HLA-B7 levels seemed to be similar in either the B7 single transgene construct or the CD2:B7 construct.

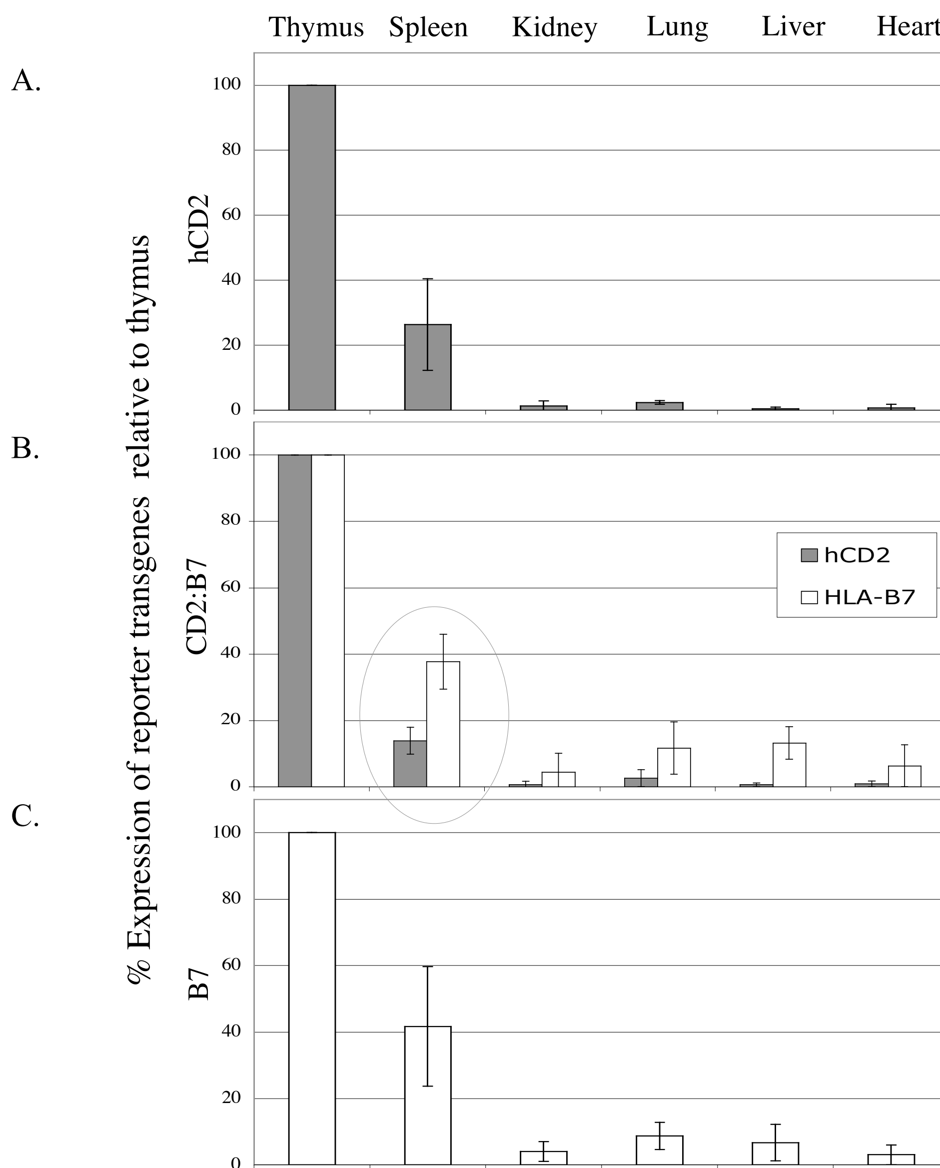


Figure 15. Pattern of expression in transgenes is similar to TCRα expression. Regardless of orientation, both reporter genes exhibited TCRα-like expression patterns. Northern blots were quantified for expression relative to thymus. Multiple tissues indicated above were extracted from three independent lines per construct. Graphs represent the average relative tissue distribution of normalized transgene expression in the indicated organs. Error bars represent the SD in percent thymic expression levels for each organ within a particular construct. Grey bars represent hCD2 expression, open bars represent HLA-B7 expression. A. Average expression in three independent hCD2 lines relative to thymus. B. Average expression in three independent CD2:B7 lines relative

to thymus. C. Average expression in three independent HLA-B7 lines relative to thymus. Note that the bars within the oval indicate a difference of only hCD2 expression, not B7 expression, in the spleen when compared to the single transgene hCD2 or HLA-B7.

4.1.4 hCD2 ectopic B cell expression is suppressed in CD2:B7 transgenes

The results described above pointed to studies that would provide a more extensive analysis of spleen expression in the double CD2:B7 transgene compared to that of the single transgene constructs. Using flow cytometry, whole thymus and spleen tissue was made into a single cell suspension. Paired thymocytes and spleen B subpopulations were compared for further analyses. Thymocytes, CD19+ splenic B cell populations and CD90+ spleen T cell populations in non-transgenic and transgenic mice were compared for hCD2 expression.

Previously reported data showed a high incidence of ectopic B cell expression in transgenes linked to the 5' portion of TCRα LCR (Harrow and Ortiz 2005). In the CD2:B7 double reporter transgene context, this ectopic B cell expression is suppressed. We compared the single hCD2 transgene to the double transgene. By flow cytometry, nearly all of the B-cells in the hCD2 lines ectopically expressed the reporter transgene (Harrow and Ortiz 2005). However, in the CD2:B7 lines, only a very small subset showed any expression of hCD2 (Figure 16).

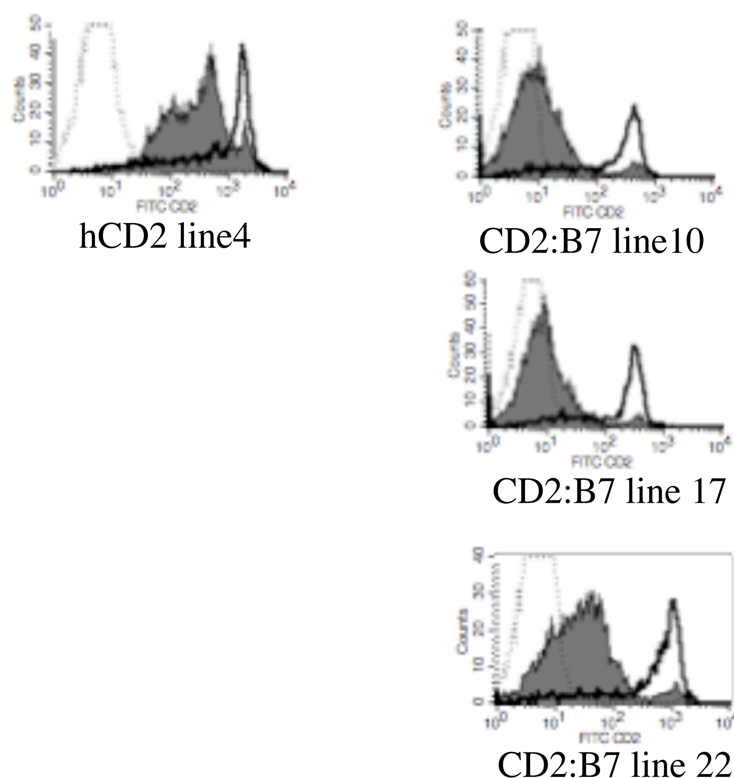


Figure 16. Ectopic expression of upstream TCRα LCR-linked reporter gene (hCD2) is suppressed by adding a downstream gene (HLA-B7) flanking the LCR. Representative flow cytometry analyses of spleen lymphocytes from (left panel) single-reporter hCD2-LCR and (right panels) two-reporter construct CD2:B7 transgenic mice. hCD2 levels are shown for CD90 positive spleen T cells (solid black line), CD19 positive spleen B cells (grey area), and non-transgenic spleen cells (dotted line). Data collected by Blanca Andino.

Relative to the thymus, the mean fluorescence intensity (MFI) of hCD2 reporter expression in B cells is very low when the LCR is in the presence of an additional downstream reporter gene. hCD2 expression in the single hCD2 construct relative to thymus level was consistent with previous results (Harrow and Ortiz 2005) ranging from 15-41% while the double CD2:B7 transgene expression ranges from 2-9% (Figure 17).

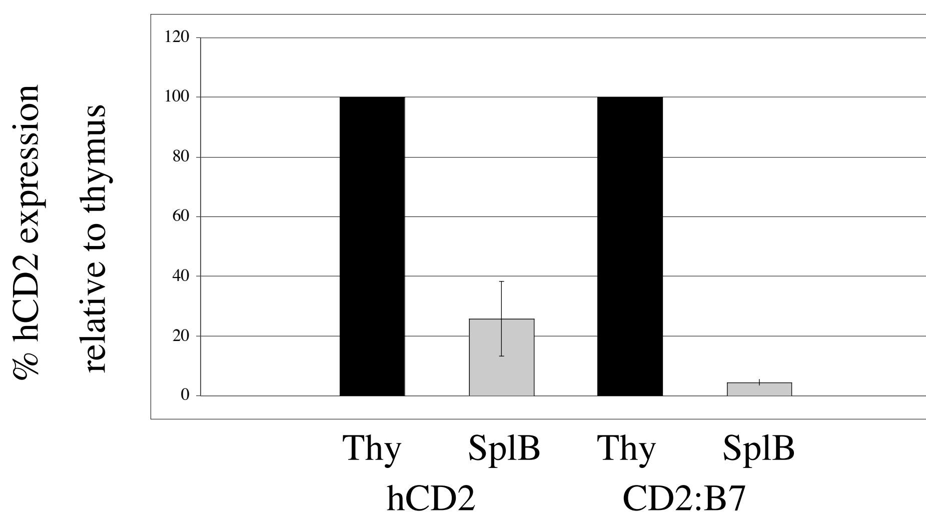
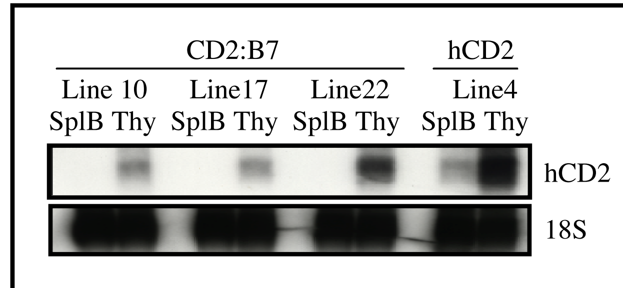


Figure 17. Mean fluorescence intensity (MFI) of hCD2 reporter staining in paired samples of whole thymocytes and spleen B cells (from the same mouse). Shown are averages from 2 independent lines of the single reporter transgene, hCD2, and three independent lines of the CD2:B7 double reporter transgenic mice (error bars=SD). MFI was normalized to thymocyte hCD2 expression (100%). In all cases, substantial ectopic B cell expression of hCD2 from the single reporter transgene was lost in the presence of a second reporter gene downstream of the LCR. Data collected by Blanca Andino.

These results were confirmed by northern blot analysis where hCD2 expression ranged from 15-38% in the single reporter hCD2 transgenic lines while expression is reduced significantly to 2-6% in the three dual reporter CD2:B7 independent lines tested (Figure 18).

A.



B.

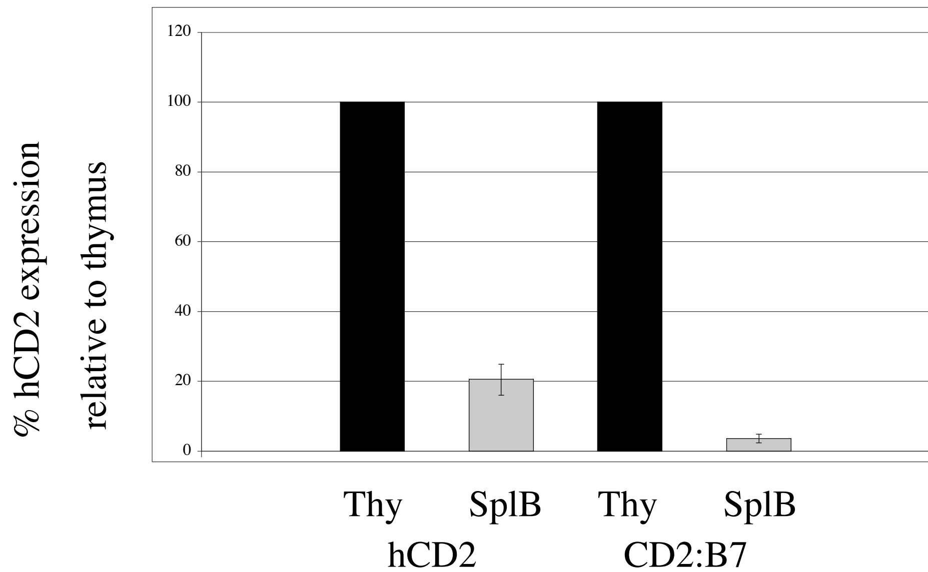


Figure 18. Ectopic hCD2 RNA levels are suppressed in CD2:B7 transgenic lines. A. Representative northern blot and B. compilation of hCD2 and CD2:B7 transgenic lines showing average hCD2 expression in Spleen B cells relative to thymus. Data points collected by Karl Erhard.

4.1.5 HLAB7 Ectopic B cell expression persists in transgenes placed 3' of the TCRα LCR.

We then compared HLA-B7 reporter transgene expression in Spleen B cells. Similar to the method above, HLA-B7 expression was compared in the CD2:B7 and B7 transgenic lines by northern blot analysis. For each of the lines, expression was compared relative to thymus.

Consistent with the tissue distribution results, we found that HLAB7 was expressed at similar ectopic levels in the double CD2:B7 transgene and single B7 transgene constructs. B7 expression relative to thymus in B7 transgenic spleen B cells ranged from 27-37% while expression of B7 from the CD2:B7 transgene maintained similar levels of expression at 18-33% (Figure 19). These results are similar to the single hCD2 transgene described above where ectopic expression of its reporter gene was 15-38%.

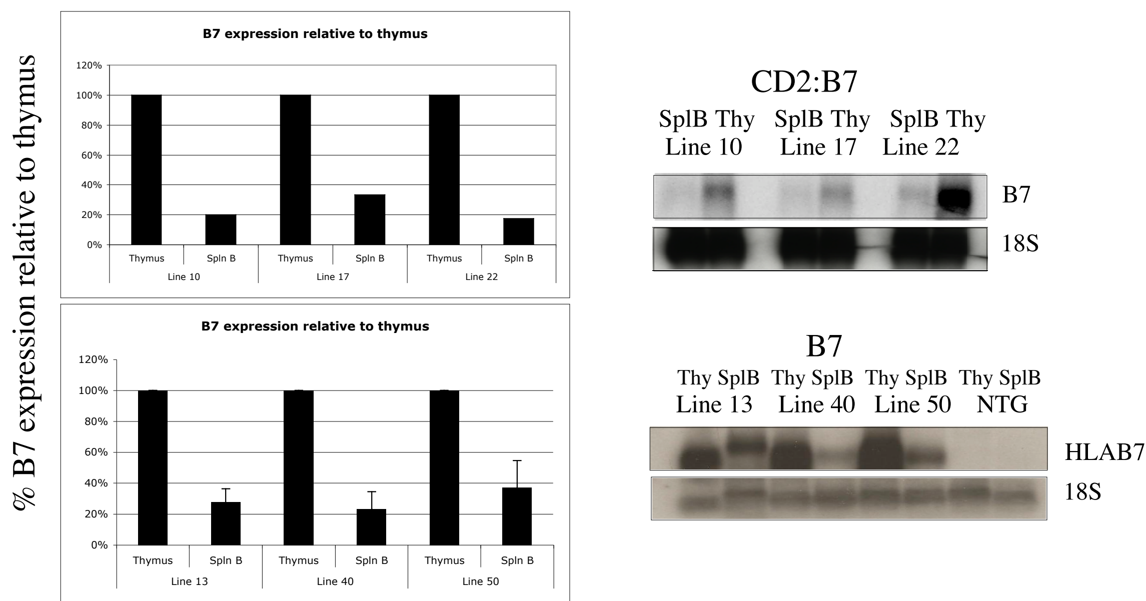


Figure 19. Spleen B cell expression (relative to thymus) of the HLA-B7 reporter is unchanged by the addition of a second reporter gene upstream of the LCR. PhosporImager analyses of northern blots of HLA-B7 comparing expression of purified spleen B cells relative to thymic expression (from the same mouse). Shown are six independent lines of either the single B7 reporter transgene construct (top panels) or the two reporter CD2:B7 (bottom panels) construct in transgenic mice. HLA-B7 mRNA levels were normalized to 18S RNA as a loading control.

4.1.6 TCRα LCR cannot maintain hCD2 Peripheral T cell expression

In the endogenous locus, TCRα expression increases in the periphery to over 50% higher than that of thymic expression. Previous data showed that that the TCRα LCR does not support high level of peripheral T cell expression in the single hCD2 transgene constructs (Harrow and Ortiz 2005). We compared peripheral reporter gene expression in CD2:B7 with that of the previously published hCD2 constructs using flow cytometry. The MFI of hCD2 in the peripheral CD90 positive cells spleen cells were measured. T cell expression

was compared to that of the thymus. Expression relative to thymus remained at the same reduced levels at an average of 42% (Figure 20). Thus, in both the single and double transgene context, the TCRα LCR is still unable to support the upregulated expression that is seen in the endogenous locus.

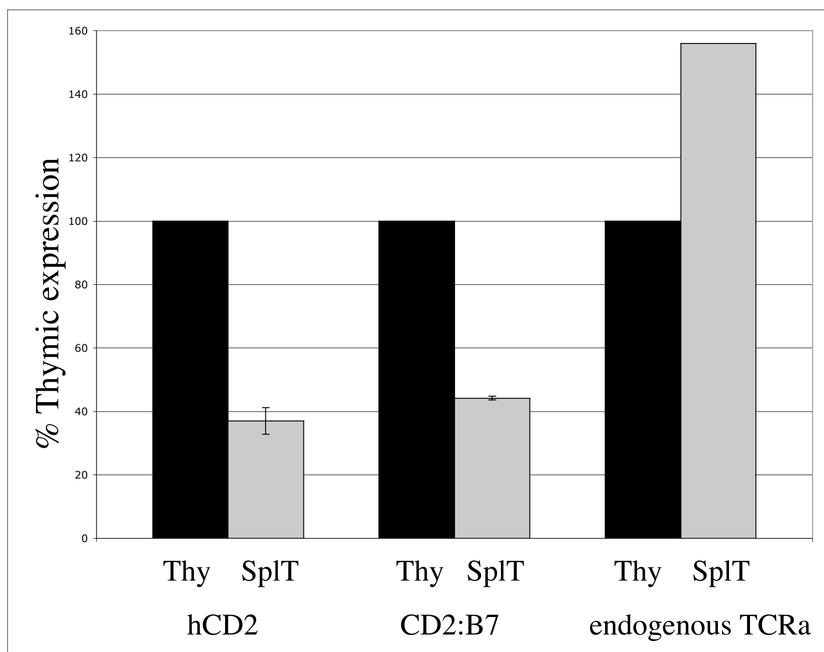


Figure 20. Peripheral T cell expression is not maintained in the double CD2:B7 transgenic lines. Flow cytometric analysis of hCD2 expression in thymocytes and spleen T cells. The relative mean fluorescence of hCD2 transgene expression is shown graphically relative to thymocytes. As a reference, previous data by northern blot analysis shows endogenous TCRα expression is increased to over 50% of thymus (Harrow and Ortiz 2005). Data collected by Blanca Andino.

4.2 SIGNIFICANCE

4.2.1 TCRα LCR activity is orientation independent

Here, we report that the TCRα LCR is orientation independent with respect to both its ability to drive copy-number related transgene expression in a TCRα -like pattern. In the transgenes containing single and dual reporter LCR-driven construct, both reporter genes displayed highest expression in lymphoid organs. We found this especially interesting given that Dad1, the gene normally located 3' to this LCR, is a ubiquitously active gene.

It has been reported that the human β-globin LCR is orientation dependent. An inverted LCR is incapable of activating downstream globin genes at a high level (Tanimoto et al. 1999; Routledge and Proudfoot 2002). In addition, a gene placed upstream of the β-globin LCR was expressed at substantially lower levels than that of the endogenous gene 65kb downstream of it (Wai et al. 2003). When this LCR was placed in a gene rich locus, changes in gene expression occurred bi-directionally and as far as 150 kb from the site of LCR insertion, but were variable for individual genes, depending on the orientation of the integrated LCR (Noordermeer et al. 2008).

However, other LCRs are orientation independent. The mouse tyrosinase LCR is able to overcome position effects in cell culture regardless of orientation (Giraldo et al. 2003). In addition, the human growth hormone (hGH) LCR was shown to enhance expression in cell lines and transgenic studies in both the forward and reverse orientation (Gimenez et

al. 2001).

Because the TCRα LCR lies closely between 2 differentially regulated genes, we were surprised to see an orientation independent function of the TCRα LCR. It is apparent from our data that the TCRα LCR is important in maintaining TCRα expression patterns. However, the endogenous locus may offer other regulatory mechanisms to influence the directionality and quality of LCR activity.

4.2.2 LCR activity is influenced differently when flanked by separate reporter genes

Our results indicate that the LCR can participate in the regulation of genes placed 5' or 3' of it. In single reporter transgenes, regardless of orientation, LCR activity is similar. This indicates that the LCR might adopt a 'default' activating state whereby a single gene linked to the LCR is expressed in a consistent TCRα-like pattern regardless of its position with respect to the LCR. However, the endogenous gene 3' of the TCRα LCR is Dad1, whose expression is ubiquitous. Therefore, other Dad1 elements outside of the LCR must act to upregulate or modulate its expression to normal physiological levels in non-lymphoid tissues.

In the dual reporter gene context, we further found that ectopic B cell expression of the upstream gene was lost while T cell expression still remained the same relative to the

single hCD2 transgene. This provides evidence that a bi-directional LCR activity we discovered is asymmetric in quality. From our data, it seems that only the 3' directed regulatory activity exists in non-T cells. This is because in B cells, we observed that the downstream gene is activated while the upstream gene is not. In contrast, both the 5'- and 3'- directed regulatory activity exist in T cells allowing efficient co-expression of both genes. The apparent separation of the 3' directed, non-T cell specific regulation from the 5' directed T cell specific regulatory activity seen in B cells (and perhaps all non-lymphoid cells) could be enforced by an insulator and/or supported by factors exclusive to particular cell types. Consistent with what is known about the TCRα LCR, multiple activating and insulator-like regions have been identified. One of these regions, such as the CTCF binding sites in HS1', could participate in the separation of LCR activity when faced with two actively regulated genes on either side of it (Figure 21).

The CTCF binding site is a proven enhancer-blocking sequence in the TCRα LCR (Magdinier, Yusufzai, and Felsenfeld 2004) and could play a role in separating the LCR's two regulatory activities. In other loci, evidence shows that part of CTCF's functionality depend on cell specific factors. It has been shown that CTCF must interact with at least two additional factors, RFX and CIITA to function in the control of the *HLA-DRB1* and *HLA–DQA1* genes (Majumder et al. 2008). The loss of these factors abolished long-range interactions with the XL9 enhancer element while the induction of CIITA, an immune cell specific gene, in epithelial cells caused CTCF recruitment. Similarly, T cells may block CTCF from negatively influencing the LCR in a cell specific manner. The ectopic B cell expression would be lost in the dual reporter transgene context because CTCF

would be able to form a complex or barrier that prevents upstream interaction. This action that prevents upstream gene activation in specific cells can be seen by the lack of ectopic B cell reporter gene expression in only the upstream gene in a dual reporter context. This function would not have been revealed in previous single reporter transgenic studies. This is because the transgenic lines studied contained multiple copies of the transgene in head-to-tail tandem arrays. Therefore both the 3' and 5'-directed regulatory activities of the LCR act together on the same reporter gene. This would produce the largely T cell specific pattern (with ectopic B cell expression) previously observed. By comparing LCR activity in single and double reporter transgene contexts, we were able to discern that the LCR indeed has bi-directional regulatory capabilities. Furthermore, the 5' directed activity is strictly T cell specific. In contrast, the 3' directed activity while still highly active in T cells, is less tissue restricted.

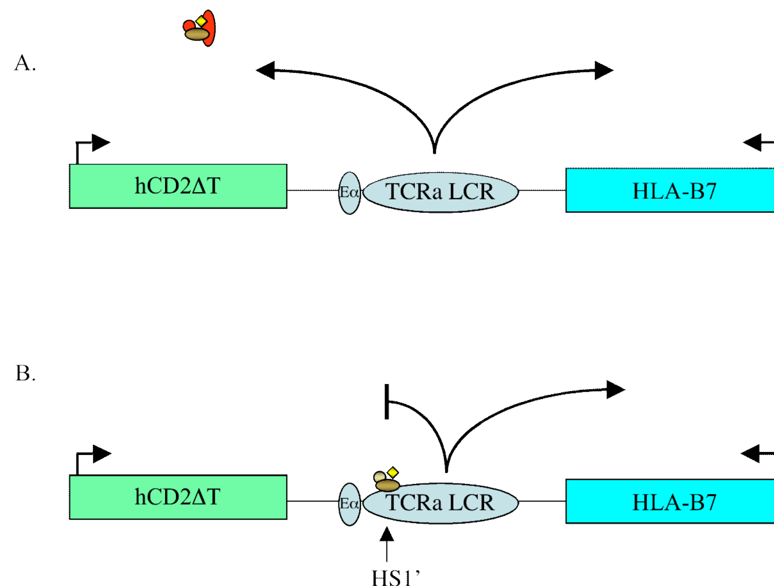


Figure 21. Model of bi-directional LCR gene expression regulation. Green or teal boxes indicate reporter genes, hCD2 and B7, and blue ovals represent the TCRα LCR. Horizontal arrows show transcription orientation. Activator function is indicated by the curved line with arrow while

inhibition/repression is indicated by the vertical line. Small shape clusters represent factors that could bind CTCF to prevent or to allow active binding in HS1'. A. In T-cells, the TCRα LCR activates a level of expression that is lymphoid specific in regardless of gene placement relative to the LCR. T cell specific factors can interfere with CTCF's ability to bind at HS1' B. In non-T cells, where HS1' is predominant and CTCF sites exist, other factors allow CTCF to bind HS1'. This interaction causes insulator activity from some of the activating regions of the LCR preventing expression of the upstream hCD2 gene.

4.3 FUTURE DIRECTION

The molecular mechanism for LCR activity in the dual reporter transgene context is still unknown. Part of our on-going studies are to look at preferential promoter accessibility or permissive chromatin modifications in spleen B cells in the varying single and double transgene constructs that have been analyzed.

4.3.1 DNase I Hypersensitivity assays (DHA)

DNase1 hypersensitivity assays (DHA) can be used to compare chromatin structure. B cells from both the single hCD2 and double CD2:B7 transgenic lines can be isolated to pursue these studies. Because ectopic B cell expression is suppressed in the dual reporter transgene, we would expect to see differences in chromatin structure at the reporter gene promoter region or at the LCR itself. This would indicate a specific effect due to the presence of the downstream gene in the dual reporter system. If hypersensitivity patterns were unchanged between the two, the level of regulation may be by differential factor interaction or at a higher level of chromatin modification.

4.3.2 Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) can be done to show differences that affect gene activity in a chromatin environment. Many studies show that transcription is also regulated by 'the histone code'. Here, epigenetic post-translational modifications like acetylation and methylation markers on histone tails can indicate what conformation a discrete chromatin domain will be in. For example, modified lysine and arginine residues can exist in several methylated states. In addition, ChIP technology can be further utilized to identify transcription factor binding in very focused regions.

One particular use of ChIP to identify mechanisms of LCR activity in the dual reporter transgene context is to compare the B cell population of the single hCD2 reporter and double CD2:B7 reporter transgenes. The reporter genes and specific regions of the LCR would be screened for active histone markers like tri-methylated lysine 4 on histone H3. From these types of experiments, I would expect to see differences in histone modification or in factor binding that would represent how the genes are differentially suppressed in the CD2:B7 context.

CHAPTER 5: Creating a BAC reporter transgene model to study TCRα/Dad1 regulation

5.1 Results

The transgenic models for the TCRα LCR have not been able to replicate all of the regulatory aspects seen in the endogenous locus. Specifically, there is a lack of upregulated expression of TCRα in peripheral T cells and the inability for the transgene to specify Dad1 ubiquitous expression patterns ((Harrow and Ortiz 2005),data described above). Our findings indicate that other elements within the locus are needed or may regulate these specific functions.

More aspects of LCR activity and regulation within the locus can be investigated by creating a transgenic model that contains a larger portion of the endogenous locus. This would enable us to identify new regulatory elements within the locus that drive TCRα and Dad1 expression. This can be accomplished using a modified bacterial artificial chromosome (BAC) reporter system. BAC vectors can accommodate large genomic intervals and are valuable for characterizing genes and elements, like a locus control region, that contribute to differential expression over large regions of a gene locus. This reporter construct can be manipulated to analyze the consequences of various deletions or mutations within the LCR for reporter gene expression.

5.1.1 Bacterial Artificial Chromosomes

Bacterial artificial chromosomes (BAC) contain large segments of DNA loci and have been used in transgenic studies. Two advantages in using a BAC for research purposes are the relative ease with which DNA can be modified and the ability to focus research in a more endogenous, but definable, environment. BACs can harbor up to 300 kb, containing several contiguous genes along with their cis-acting regulatory elements. This provides enough sequence to allow for the natural spacing and identification of novel control elements in a moderately simple construct.

5.1.2 BAC containing the TCRα LCR

The TCRα locus is over 1.5Mb. However, much of that size is due to the V and J gene segments upstream of the TCRα LCR. The original BAC clone used in this study is RP23-94I14 (I14) from the BACPAC Resources Center (BPRC) containing 46105796-46287953 base pairs from mouse Chromosome 14. This ~180kb (182157bp) BAC contains the TCRα locus, including TCRδ, the upstream variable J regions of TCRα, TCRα constant region, LCR, Dad1 gene, and several other downstream genes. This region is flanked by Not1 restriction sites in the pBACe3.6 vector (http://bacpac.chori.org/pbace36.htm) that can be used to excise the insert.

5.1.3 Modified TCRα LCR BAC

By modifying the above BAC, we inserted reporter genes in the TCRα and Dad1 orientation while keeping all the TCRα LCR elements and other functionally unidentified neighboring sequence intact (Figure 20). Once this model is established, we can use it to analyze the effects of specific deletions and mutations in the LCR or other identified elements.

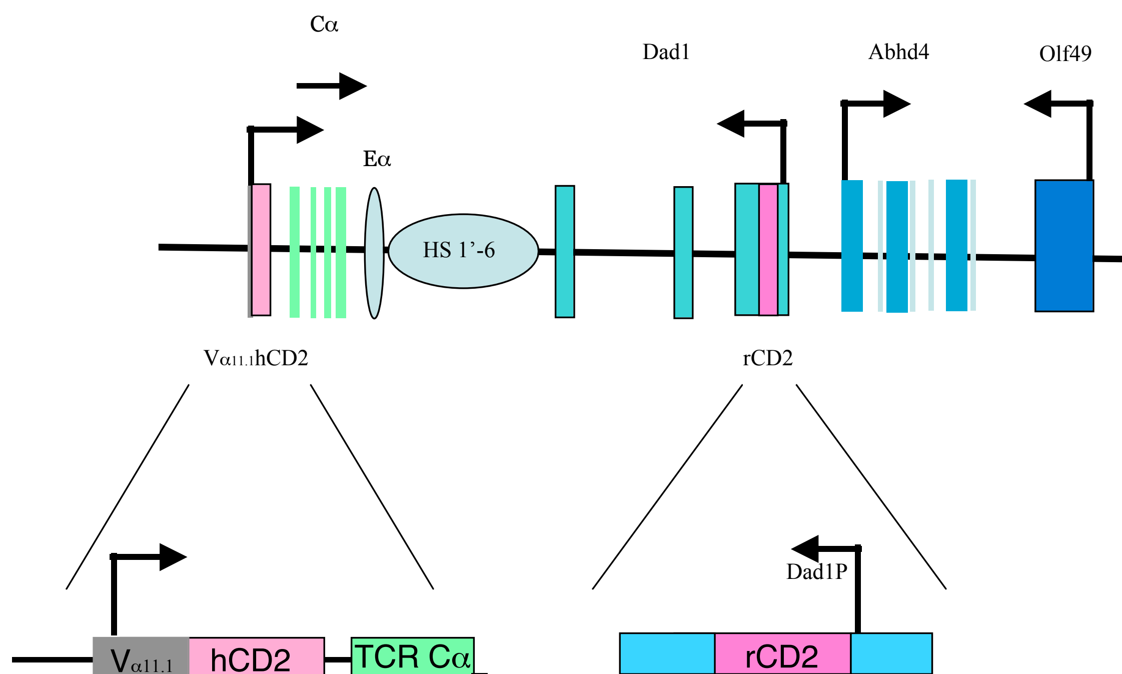


Figure 22. Diagram of the TCRα/Dad1 Bacterial artificial chromosome (BAC) reporter construct, Δ5'I14h/rCD2. It contains TCRα V-region promoter driving human cytoplasmic tail-less CD2 cDNA and Dad1 promoter driving rat cytoplasmic tail-less CD2 cDNA. The TCRα constant region exons (green boxes), HS of the TCRα LCR (ovals) and Dad1 exons (teal boxes) are shown. The Not I sites used to liberate fragment are ~80kb apart.

In order to focus on elements that affect TCRα and Dad1 gene expression, we used a BAC containing over 80kb of endogenous sequence. This sequence includes the TCRα constant region, the TCRα LCR, Dad1 and all of the natural components between them and several genes downstream of Dad1. The I14 BAC was modified in our lab to add a functional Vα11.1 region (from Vα11.1Jα14.4 of TCRα (Diaz, Cado, and Winoto 1994)) and 2 reporter genes, human and rat CD2, in the TCRα and Dad1 orientations respectively. CD2 is a 50- to 55-kDa glycoprotein cell adhesion molecule that normally mediates T-cell activation; it also acts as a cell surface marker. The TCRα Vα11.1 segment was fused in frame at the ATG with the reporter gene encoding a cDNA for cytoplasmic tail-less human CD2ΔT (hCD2ΔT). This reporter gene lacks 100 C-terminal amino acids of the cyctoplasmic domain required for downstream signaling events (Melton et al. 1996). This Vα11.1hCD2ΔT reporter construct was put in the TCRα orientation just upstream of Cα. Similarly, a 703bp rat CD2 (also cytoplamic tail-less) reporter gene (Hozumi et al. 2000) was inserted at the ATG of the first exon of Dad1. These modifications make a construct containing endogenous TCRα and Dad1 promoters driving reporter genes while still having elements that exist in a chromosomal context. The described modified BAC construct was called I14 h/rCD2.

Success in generating clones decreases with larger insert sizes and can cause increased rearrangements (Shizuya and Kouros-Mehr 2001). Furthermore, insertion of fragmented transgenes can occur with an increase in transgene size. Thus, the 5' 90,341 base pairs of the BAC which contains TCRδ and TCRα variable J regions were deleted in order to make a smaller, more manageable ~90kb BAC. This BAC, called Δ5'I14h/rCD2, is the

core construct used for these studies.

5.1.4 Confirm modified BAC sequence

The promoter region, reporter genes, and the LCR were sequenced to confirm that the construct is as expected. Primers made at both Vα11.1 and Dad1 promoters verified that the reporter genes are in-frame to its endogenous gene's ATG start site. Because the reporter gene fragments were amplified using PCR, I also confirmed that the cDNA matched its published sequence.

5.1.5 Isolating BAC fragment for transgenic injection

Following confirmation of Δ5'I14 human/rat CD2, the insert was excised by Not1 and purified for injection as described. Analysis was done by southern blot and PCR to ensure that there was no vector contamination and that both reporter genes were present. To genotype founders, primers were designed to the Vα11.1 promoter region and hCD2 cDNA. At least 3 founder lines have been established and are still being analyzed.

5.2 SIGNIFICANCE

5.2.1 TCRα reporter gene expression

As mentioned above, the TCRα LCR does not participate in the upregulation or maintenance of TCRα reporter genes in the periphery. A potential reason that this aspect of endogenous TCRα gene expression was not observed in TCRα LCR-driven transgenes is that it is dependent on yet-to-be identified cis-elements in the wider TCRα /Dad1 genomic locus. By looking at TCRα expression in a more endogenous like environment, we hope to identify specific elements needed to control this separate and important function.

Our findings that placement of a gene downstream of TCRα may induce silencing of an ectopic (i.e. B cell) upstream gene expression indicates that the regulation of TCRα expression must be considered in a larger, more endogenous locus like environment. Interestingly, our data suggests that Dad1 expression downstream of the LCR may help impose a T cell restriction on the LCR's activity on the upstream TCRα gene.

We hoped to study the expression of both reporter genes by flow cytometry or northern blot analyses. Using methods previously described, we also want to look for ectopic B cell expression and expression in the periphery for the TCRα oriented reporter gene. We expected to see expression of human and rat CD2 at similar levels to the endogenous

TCRα and Dad1 genes, respectively. If the modified BAC contained all the regulatory elements needed for expression at an endogenous level, we expect that rCD2 (reporter in Dad1 orientation) should be constitutively expressed while hCD2ΔT (reporter in TCRα orientation) would be expressed in only lymphoid tissues.

Our preliminary results show that the hCD2 reporter transgene is not expressed at a detectable level. This may be due to the low activity of the Vα promoter that was inserted, a lack of exon/intron processing needed for expression or a lack of poly-A signal at the end of the reporter cDNA. The Vα promoter that was used was previously used in other transgenic contexts (Diaz, Cado, and Winoto 1994). In addition, we used a previously described cDNA reporter gene. However, some studies have shown that exon/intron splicing need to occur for some genes to be properly expressed. In some cases, poly-A signaling has been shown to be necessary for genes to be expressed efficiently in order to avoid nonsense mediated decay (reviewed in (Baker and Parker 2004)). We are currently designing a new BAC vector that addresses these issues. However, because the transgene in the Dad1 orientation can be detected, this transgenic model may give us insight to Dad1 regulatory influences.

5.2.2 Dad1 reporter gene expression

What maintains the very separate expression profile for Dad1 has yet to be determined. Our observations thus far have led us to ask: What aspects of the endogenous gene

expression patterns for the TCRα /Dad1 can be reproduced by the TCRα LCR? In addition, our studies with the "double transgene" transgenic lines described above indicated that the Dad1 expression profile rely on other not yet identified factors. This BAC can provide a way to identify elements that contribute to Dad1 expression.

5.3 FUTURE DIRECTION

We have already successfully obtained multiple lines of transgenic mice bearing the Δ5'I14h/rC2 BAC fragment. However, characterization is still underway.

5.3.1 Copy number of transgenic lines

Copy number will be confirmed, although BACs tend to integrate in a single copy fashion. Probes for each reporter gene have been made from plasmids containing cDNA of human CD2ΔT (pMIh-CD2 (Deftos et al. 1998)) and rat CD2 (pMXr-CD2 (Hozumi et al. 2000)). These probes will be hybridized with genomic DNA digested with Sac I from Δ5'I14 h/r CD2 founders. The expected fragments will be 2.5kb and 4.65kb for human CD2ΔT and rat CD2 respectively.

5.3.2 Expression of reporter transgenes

Although hCD2 cannot be detected in this described BAC construct, the reporter gene in

the Dad1 orientation may provide us with a model for what elements are required in its regulation. It has been suggested that because of its location, Dad1 may require a dynamic type of regulation that prevents it from activating tightly regulated TCRα gene expression in non-T cells yet allows upregulated expression during stages of T-cell development.

5.3.3 BAC model system

Using the above methods, a BAC dual reporter transgene construct and derivatives of it can be designed for further functional studies. Transgenic mice can be a costly and long term endeavor. Because BACs are relatively easy to modify, this model system can be a core from which mutant constructs can be used in transfected cell lines or stem cells. A system such as this can give insight to LCR involvement, or other functional sequences, in the transcriptional control of genes in the TCRα and Dad1 orientations at a whole locus level.

CHAPTER 6: Discussion

Overall, the data that I've presented here point to several novel findings. These findings give insight to LCR regulation that may ultimately help in the design for better gene therapy vectors that require predictable amounts of tissue specific expression in lymphoid tissues.

From our data, I proposed a potential model for how TCRα LCR activation and regulation affect its differentially expressed neighboring genes (Figure 21). In this model, the LCR has an activating influence on both its neighboring genes. However, this activation can be further controlled by the CTCF binding sites that exist in HS1'. In T cells, cell specific factors can prevent CTCF from interacting productively with HS1'. This action allows the activating signals from the entire LCR to influence its neighboring genes (upstream and downstream). In non-T cells, factors are able to recruit and/or activate the binding of CTCF to HS1'. This would cause HS1' to act as an insulator element in non-T cells and prevent LCR activation upstream of HS1' while still allowing LCR activation at the downstream gene. This type of regulation would cause an increased tissue specificity of the upstream gene in non-T cells while still allowing overall expression of the downstream gene. The TCRα LCR confers a lymphoid tissue specific expression profile to genes regardless of orientation. Dad1 is ubiquitously expressed and located just downstream of the LCR. Elements within the Dad1 region (or downstream gene) would need to upregulate and modulate expression in non-lymphoid tissues. This

model could also explain why HS1', where the CTCF binding sites are, is predominant in non-lymphoid tissues.

In addition, the TCRα LCR contains other multiple insulator-like regions that are non-redundant. These elements might contribute to a dynamic regulation that gets modulated especially during different stages of T cell development. The insulator-like elements could differentially interact with each other or activating regions within the locus so that expression can be tightly regulated in T cell development or cell type. In this model, the TCR gene is either poised for expression or kept inactive with the help of the insulator elements. Depending on the stage of T cell development or cell type, TCR expression must also be shielded from the additional activating effects from Dad1 regulatory elements. For example HS4 might shield from the effects of ubiquitous Dad1 expression. A deletion of HS4 in the transgenic context support this since, without HS4, the reporter genes are not protected from position effect and expression is no longer controlled to a specific tissue. HS4 is also specifically demethylated in lymphoid tissues. Although its function is still unknown, this may indicate that HS4 is utilized specifically in lymphoid tissues.

Because the TCRα LCR has the powerful ability to overcome position effect and express transgenes at specific and predictable levels, it can be an ideal regulatory element to use in specific therapeutic applications, such as gene therapy. Our data shows that the LCR can activate expression in a bidirectional manner. Thus, if the therapeutic genes are in concatamers or placed 3' of the LCR, some ectopic expression may be detected. This

aspect, especially if the therapeutic gene must be tightly regulated and exclusive to T cells, must be taken into account. In order to control for this, the LCR linked therapeutic gene may need to be placed in either a single copy or with additional insulators flanking it so that each transcription unit can be an isolated as a distinct regulatory gene unit. Another problem in designing gene therapy vectors is the size of insert that is put into its delivery vehicles. To cut back on size, a "mini-LCR" containing only the essential sub elements could be designed. In this way, the work presented here identifies and characterizes key functional LCR sub-elements that will contribute to effort of applying aspects of LCR activity to gene therapy.

REFERENCES

Baker, K. E., and R. Parker. 2004. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol* 16 (3):293-9.

Baldwin, T. A., K. A. Hogquist, and S. C. Jameson. 2004. The fourth way? Harnessing aggressive tendencies in the thymus. *J Immunol* 173 (11):6515-20.

Bell, A. C., A. G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98 (3):387-96.

Blackwood, E. M., and J. T. Kadonaga. 1998. Going the distance: a current view of enhancer action. *Science* 281 (5373):60-3.

Brenner, M. B., J. L. Strominger, and M. S. Krangel. 1988. The gamma delta T cell receptor. *Adv Immunol* 43:133-92.

Bulger, M., and M. Groudine. 1999. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev* 13 (19):2465-77.

Carter, D., L. Chakalova, C. S. Osborne, Y. F. Dai, and P. Fraser. 2002. Long-range chromatin regulatory interactions in vivo. *Nat Genet* 32 (4):623-6.

Chamberlain, J. W., J. A. Nolan, P. J. Conrad, H. A. Vasavada, H. H. Vasavada, H. L. Ploegh, S. Ganguly, C. A. Janeway, Jr., and S. M. Weissman. 1988. Tissue-specific and cell surface expression of human major histocompatibility complex class I heavy (HLA-B7) and light (beta 2-microglobulin) chain genes in transgenic mice. *Proc Natl Acad Sci U S A* 85 (20):7690-4.

Chang, H. C., P. Moingeon, P. Lopez, H. Krasnow, C. Stebbins, and E. L. Reinherz. 1989. Dissection of the human CD2 intracellular domain. Identification of a segment required for signal transduction and interleukin 2 production. *J Exp Med* 169 (6):2073-83.

Chauveau, C., E. A. Jansson, S. Muller, M. Cogne, and S. Pettersson. 1999. Cutting edge: Ig heavy chain 3' HS1-4 directs correct spatial position-independent expression of a linked transgene to B lineage cells. *J Immunol* 163 (9):4637-41.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162 (1):156-9.

de Laat, W., and F. Grosveld. 2003. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* 11 (5):447-59.

Defossez, P. A., and E. Gilson. 2002. The vertebrate protein CTCF functions as an insulator in Saccharomyces cerevisiae. *Nucleic Acids Res* 30 (23):5136-41.

Deftos, M. L., Y. W. He, E. W. Ojala, and M. J. Bevan. 1998. Correlating notch signaling with thymocyte maturation. *Immunity* 9 (6):777-86.

Diaz, P., D. Cado, and A. Winoto. 1994. A locus control region in the T cell receptor alpha/delta locus. *Immunity* 1 (3):207-17.

Dorsett, D. 1999. Distant liaisons: long-range enhancer-promoter interactions in Drosophila. *Curr Opin Genet Dev* 9 (5):505-14.

Dudley, E. C., H. T. Petrie, L. M. Shah, M. J. Owen, and A. C. Hayday. 1994. T cell receptor beta chain gene rearrangement and selection during thymocyte development in adult mice. *Immunity* 1 (2):83-93.

Engel, J. D., and K. Tanimoto. 2000. Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. *Cell* 100 (5):499-502.

Enver, T., A. C. Brewer, and R. K. Patient. 1985. Simian virus 40-mediated cis induction of the Xenopus beta-globin DNase I hypersensitive site. *Nature* 318 (6047):680-3.

Fehling, H. J., A. Krotkova, C. Saint-Ruf, and H. von Boehmer. 1995. Crucial role of the pre-T-cell receptor alpha gene in development of alpha beta but not gamma delta T cells. *Nature* 375 (6534):795-8.

Felsenfeld, G., B. Burgess-Beusse, C. Farrell, M. Gaszner, R. Ghirlando, S. Huang, C. Jin, M. Litt, F. Magdinier, V. Mutskov, Y. Nakatani, H. Tagami, A. West, and T. Yusufzai. 2004. Chromatin boundaries and chromatin domains. *Cold Spring Harb Symp Quant Biol* 69:245-50.

Festenstein, R., M. Tolaini, P. Corbella, C. Mamalaki, J. Parrington, M. Fox, A. Miliou, M. Jones, and D. Kioussis. 1996. Locus control region function and heterochromatin-induced position effect variegation. *Science* 271 (5252):1123-5.

Filippova, G. N., M. K. Cheng, J. M. Moore, J. P. Truong, Y. J. Hu, D. K. Nguyen, K. D. Tsuchiya, and C. M. Disteche. 2005. Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Dev Cell* 8 (1):31-42.

Gaszner, M., and G. Felsenfeld. 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7 (9):703-13.

Gimenez, E., P. Giraldo, G. Jeffery, and L. Montoliu. 2001. Variegated expression and delayed retinal pigmentation during development in transgenic mice with a deletion in the locus control region of the tyrosinase gene. *Genesis* 30 (1):21-5.

Giraldo, P., A. Martinez, L. Regales, A. Lavado, A. Garcia-Diaz, A. Alonso, A. Busturia, and L. Montoliu. 2003. Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. *Nucleic Acids Res* 31 (21):6290-305.

Godfrey, D. I., J. Kennedy, P. Mombaerts, S. Tonegawa, and A. Zlotnik. 1994. Onset of TCR-beta gene rearrangement and role of TCR-beta expression during CD3-CD4-CD8- thymocyte differentiation. *J Immunol* 152 (10):4783-92.

Godfrey, D. I., J. Kennedy, T. Suda, and A. Zlotnik. 1993. A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. *J Immunol* 150 (10):4244-52.

Goldrath, A. W., and M. J. Bevan. 1999. Selecting and maintaining a diverse T-cell repertoire. *Nature* 402 (6759):255-62.

Gomos-Klein, J., F. Harrow, J. Alarcon, and B. D. Ortiz. 2007. CTCF-independent, but not CTCF-dependent, elements significantly contribute to TCR-alpha locus control region activity. *J Immunol* 179 (2):1088-95.

Greaves, D. R., F. D. Wilson, G. Lang, and D. Kioussis. 1989. Human CD2 3'-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell* 56 (6):979-86.

Groettrup, M., K. Ungewiss, O. Azogui, R. Palacios, M. J. Owen, A. C. Hayday, and H. von Boehmer. 1993. A novel disulfide-linked heterodimer on pre-T cells consists of the T cell receptor beta chain and a 33 kd glycoprotein. *Cell* 75 (2):283-94.

Grosveld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51 (6):975-85.

Harju, S., K. J. McQueen, and K. R. Peterson. 2002. Chromatin structure and control of beta-like globin gene switching. *Exp Biol Med (Maywood)* 227 (9):683-700.

Harrow, F., J. U. Amuta, S. R. Hutchinson, F. Akwaa, and B. D. Ortiz. 2004. Factors binding a non-classical Cis-element prevent heterochromatin effects on locus control region activity. *J Biol Chem* 279 (17):17842-9.

Harrow, F., and B. D. Ortiz. 2005. The TCRalpha locus control region specifies thymic, but not peripheral, patterns of TCRalpha gene expression. *J Immunol* 175 (10):6659-67.

Hernandez-Hoyos, G., M. K. Anderson, C. Wang, E. V. Rothenberg, and J. Alberola-Ila. 2003. GATA-3 expression is controlled by TCR signals and regulates CD4/CD8 differentiation. *Immunity* 19 (1):83-94.

Hernandez-Munain, C., B. P. Sleckman, and M. S. Krangel. 1999. A developmental switch from TCR delta enhancer to TCR alpha enhancer function during thymocyte maturation. *Immunity* 10 (6):723-33.

Hoffman, E. S., L. Passoni, T. Crompton, T. M. Leu, D. G. Schatz, A. Koff, M. J. Owen, and A. C. Hayday. 1996. Productive T-cell receptor beta-chain gene rearrangement: coincident regulation of cell cycle and clonality during development in vivo. *Genes Dev* 10 (8):948-62.

Hong, N. A., D. Cado, J. Mitchell, B. D. Ortiz, S. N. Hsieh, and A. Winoto. 1997. A targeted mutation at the T-cell receptor alpha/delta locus impairs T-cell

development and reveals the presence of the nearby antiapoptosis gene Dad1. *Mol Cell Biol* 17 (4):2151-7.

Hong, N. A., M. Flannery, S. N. Hsieh, D. Cado, R. Pedersen, and A. Winoto. 2000. Mice lacking Dad1, the defender against apoptotic death-1, express abnormal N-linked glycoproteins and undergo increased embryonic apoptosis. *Dev Biol* 220 (1):76-84.

Hong, N. A., N. H. Kabra, S. N. Hsieh, D. Cado, and A. Winoto. 1999. In vivo overexpression of Dad1, the defender against apoptotic death-1, enhances T cell proliferation but does not protect against apoptosis. *J Immunol* 163 (4):1888-93.

Hozumi, K., R. Ohtsuka, D. Suzuki, K. Ando, M. Ito, T. Nishimura, M. Merkenschlager, and S. Habu. 2000. Establishment of efficient reaggregation culture system for gene transfection into immature T cells by retroviral vectors. *Immunol Lett* 71 (1):61-6.

Kang, J., and D. H. Raulet. 1997. Events that regulate differentiation of alpha beta TCR+ and gamma delta TCR+ T cells from a common precursor. *Semin Immunol* 9 (3):171-9.

Kelleher, D. J., and R. Gilmore. 1997. DAD1, the defender against apoptotic cell death, is a subunit of the mammalian oligosaccharyltransferase. *Proc Natl Acad Sci U S A* 94 (10):4994-9.

Kent, W. J. 2002. BLAT--the BLAST-like alignment tool. *Genome Res* 12 (4):656-64.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler. 2002. The human genome browser at UCSC. *Genome Res* 12 (6):996-1006.

Kioussis, D., E. Vanin, T. deLange, R. A. Flavell, and F. G. Grosveld. 1983. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature* 306 (5944):662-6.

Koop, B. F., L. Rowen, K. Wang, C. L. Kuo, D. Seto, J. A. Lenstra, S. Howard, W. Shan, P. Deshpande, and L. Hood. 1994. The human T-cell receptor TCRAC/TCRDC (C alpha/C delta) region: organization, sequence, and evolution of 97.6 kb of DNA. *Genomics* 19 (3):478-93.

Krangel, M. S., J. Carabana, I. Abbarategui, R. Schlimgen, and A. Hawwari. 2004. Enforcing order within a complex locus: current perspectives on the control of V(D)J recombination at the murine T-cell receptor alpha/delta locus. *Immunol Rev* 200:224-32.

Krawczyk, M., N. Peyraud, N. Rybtsova, K. Masternak, P. Bucher, E. Barras, and W. Reith. 2004. Long distance control of MHC class II expression by multiple distal enhancers regulated by regulatory factor X complex and CIITA. *J Immunol* 173 (10):6200-10.

Kushida, M. M., A. Dey, X. L. Zhang, J. Campbell, M. Heeney, J. Carlyle, S. Ganguly, K. Ozato, H. Vasavada, and J. W. Chamberlain. 1997. A 150-base pair 5' region of the MHC class I HLA-B7 gene is sufficient to direct tissue-specific expression and locus control region activity: the alpha site determines efficient expression and in vivo occupancy at multiple cis-active sites throughout this region. *J Immunol* 159 (10):4913-29.

Lang, G., C. Mamalaki, D. Greenberg, N. Yannoutsos, and D. Kioussis. 1991. Deletion analysis of the human CD2 gene locus control region in transgenic mice. *Nucleic Acids Res* 19 (21):5851-6.

Lauzurica, P., and M. S. Krangel. 1994. Temporal and lineage-specific control of T cell receptor alpha/delta gene rearrangement by T cell receptor alpha and delta enhancers. *J Exp Med* 179 (6):1913-21.

Lee, G. R., C. G. Spilianakis, and R. A. Flavell. 2005. Hypersensitive site 7 of the TH2 locus control region is essential for expressing TH2 cytokine genes and for long-range intrachromosomal interactions. *Nat Immunol* 6 (1):42-8.

Li, Q., K. R. Peterson, X. Fang, and G. Stamatoyannopoulos. 2002. Locus control regions. *Blood* 100 (9):3077-86.

Litt, M. D., M. Simpson, M. Gaszner, C. D. Allis, and G. Felsenfeld. 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293 (5539):2453-5.

Magdinier, F., T. M. Yusufzai, and G. Felsenfeld. 2004. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *J Biol Chem* 279 (24):25381-9.

Magram, J., K. Chada, and F. Costantini. 1985. Developmental regulation of a cloned adult beta-globin gene in transgenic mice. *Nature* 315 (6017):338-40.

Majumder, P., J. A. Gomez, B. P. Chadwick, and J. M. Boss. 2008. The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J Exp Med* 205 (4):785-98.

Malissen, B., L. Ardouin, S. Y. Lin, A. Gillet, and M. Malissen. 1999. Function of the CD3 subunits of the pre-TCR and TCR complexes during T cell development. *Adv Immunol* 72:103-48.

Masternak, K., N. Peyraud, M. Krawczyk, E. Barras, and W. Reith. 2003. Chromatin remodeling and extragenic transcription at the MHC class II locus control region. *Nat Immunol* 4 (2):132-7.

Matys, V., E. Fricke, R. Geffers, E. Gossling, M. Haubrock, R. Hehl, K. Hornischer, D. Karas, A. E. Kel, O. V. Kel-Margoulis, D. U. Kloos, S. Land, B. Lewicki-Potapov, H. Michael, R. Munch, I. Reuter, S. Rotert, H. Saxel, M. Scheer, S. Thiele, and E. Wingender. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 31 (1):374-8.

Melton, E., N. Sarner, M. Torkar, P. A. van der Merwe, J. Q. Russell, R. C. Budd, C. Mamalaki, M. Tolaini, D. Kioussis, and R. Zamoyska. 1996. Transgene-encoded human CD2 acts in a dominant negative fashion to modify thymocyte selection signals in mice. *Eur J Immunol* 26 (12):2952-63.

Melton, T., and M. Stoneking. 1996. Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations. *J Forensic Sci* 41 (4):591-602.

Milot, E., J. Strouboulis, T. Trimborn, M. Wijgerde, E. de Boer, A. Langeveld, K. Tan-Un, W. Vergeer, N. Yannoutsos, F. Grosveld, and P. Fraser. 1996. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87 (1):105-14.

Moingeon, P., H. C. Chang, P. H. Sayre, L. K. Clayton, A. Alcover, P. Gardner, and E. L. Reinherz. 1989. The structural biology of CD2. *Immunol Rev* 111:111-44.

Mutskov, V. J., C. M. Farrell, P. A. Wade, A. P. Wolffe, and G. Felsenfeld. 2002. The barrier function of an insulator couples high histone acetylation levels with

specific protection of promoter DNA from methylation. *Genes Dev* 16 (12):1540-54.

Nakashima, T., T. Sekiguchi, A. Kuraoka, K. Fukushima, Y. Shibata, S. Komiyama, and T. Nishimoto. 1993. Molecular cloning of a human cDNA encoding a novel protein, DAD1, whose defect causes apoptotic cell death in hamster BHK21 cells. *Mol Cell Biol* 13 (10):6367-74.

Noordermeer, D., M. R. Branco, E. Splinter, P. Klous, W. van Ijcken, S. Swagemakers, M. Koutsourakis, P. van der Spek, A. Pombo, and W. de Laat. 2008. Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. *PLoS Genet* 4 (3):e1000016.

Ogawa, M. 1993. Differentiation and proliferation of hematopoietic stem cells. *Blood* 81 (11):2844-53.

Ohlsson, R., R. Renkawitz, and V. Lobanenkov. 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* 17 (9):520-7.

Ortiz, B. D., D. Cado, V. Chen, P. W. Diaz, and A. Winoto. 1997. Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. *Embo J* 16 (16):5037-45.

Ortiz, B. D., D. Cado, and A. Winoto. 1999. A new element within the T-cell receptor alpha locus required for tissue-specific locus control region activity. *Mol Cell Biol* 19 (3):1901-9.

Ortiz, B. D., F. Harrow, D. Cado, B. Santoso, and A. Winoto. 2001. Function and factor interactions of a locus control region element in the mouse T cell receptor-alpha/Dad1 gene locus. *J Immunol* 167 (7):3836-45.

Palmiter, R. D., and R. L. Brinster. 1986. Germ-line transformation of mice. *Annu Rev Genet* 20:465-99.

Pardoll, D. M., B. J. Fowlkes, J. A. Bluestone, A. Kruisbeek, W. L. Maloy, J. E. Coligan, and R. H. Schwartz. 1987. Differential expression of two distinct T-cell receptors during thymocyte development. *Nature* 326 (6108):79-81.

Patrinos, G. P., M. de Krom, E. de Boer, A. Langeveld, A. M. Imam, J. Strouboulis, W. de Laat, and F. G. Grosveld. 2004. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev* 18 (12):1495-509.

Pearse, M., L. Wu, M. Egerton, A. Wilson, K. Shortman, and R. Scollay. 1989. A murine early thymocyte developmental sequence is marked by transient expression of the interleukin 2 receptor. *Proc Natl Acad Sci U S A* 86 (5):1614-8.

Punt, J. A., J. L. Roberts, K. P. Kearse, and A. Singer. 1994. Stoichiometry of the T cell antigen receptor (TCR) complex: each TCR/CD3 complex contains one TCR alpha, one TCR beta, and two CD3 epsilon chains. *J Exp Med* 180 (2):587-93.

Recillas-Targa, F., M. J. Pikaart, B. Burgess-Beusse, A. C. Bell, M. D. Litt, A. G. West, M. Gaszner, and G. Felsenfeld. 2002. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A* 99 (10):6883-8.

Routledge, S. J., and N. J. Proudfoot. 2002. Definition of transcriptional promoters in the human beta globin locus control region. *J Mol Biol* 323 (4):601-11.

Saint-Ruf, C., K. Ungewiss, M. Groettrup, L. Bruno, H. J. Fehling, and H. von Boehmer. 1994. Analysis and expression of a cloned pre-T cell receptor gene. *Science* 266 (5188):1208-12.

Sandelin, A., W. W. Wasserman, and B. Lenhard. 2004. ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* 32 (Web Server issue):W249-52.

Santoso, B., B. D. Ortiz, and A. Winoto. 2000. Control of organ-specific demethylation by an element of the T-cell receptor-alpha locus control region. *J Biol Chem* 275 (3):1952-8.

Schwarz, B. A., and A. Bhandoola. 2006. Trafficking from the bone marrow to the thymus: a prerequisite for thymopoiesis. *Immunol Rev* 209:47-57.

Shizuya, H., and H. Kouros-Mehr. 2001. The development and applications of the bacterial artificial chromosome cloning system. *Keio J Med* 50 (1):26-30.

Shoemaker, J., M. Saraiva, and A. O'Garra. 2006. GATA-3 directly remodels the IL-10 locus independently of IL-4 in CD4+ T cells. *J Immunol* 176 (6):3470-9.

Sleckman, B. P. 2005. Lymphocyte antigen receptor gene assembly: multiple layers of regulation. *Immunol Res* 32 (1-3):253-8.

Sleckman, B. P., C. G. Bardon, R. Ferrini, L. Davidson, and F. W. Alt. 1997. Function of the TCR alpha enhancer in alphabeta and gammadelta T cells. *Immunity* 7 (4):505-15.

Spilianakis, C. G., and R. A. Flavell. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5 (10):1017-27.

Spilianakis, C. G., M. D. Lalioti, T. Town, G. R. Lee, and R. A. Flavell. 2005. Interchromosomal associations between alternatively expressed loci. *Nature* 435 (7042):637-45.

Splinter, E., H. Heath, J. Kooren, R. J. Palstra, P. Klous, F. Grosveld, N. Galjart, and W. de Laat. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20 (17):2349-54.

Staal, F. J., M. Roederer, L. A. Herzenberg, and L. A. Herzenberg. 1990. Intracellular thiols regulate activation of nuclear factor kappa B and transcription of human immunodeficiency virus. *Proc Natl Acad Sci U S A* 87 (24):9943-7.

Takemoto, N., K. Arai, and S. Miyatake. 2002. Cutting edge: the differential involvement of the N-finger of GATA-3 in chromatin remodeling and transactivation during Th2 development. *J Immunol* 169 (8):4103-7.

Tanimoto, K., Q. Liu, J. Bungert, and J. D. Engel. 1999. Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. *Nature* 398 (6725):344-8.

Teferedegne, B., M. R. Green, Z. Guo, and J. M. Boss. 2006. Mechanism of action of a distal NF-kappaB-dependent enhancer. *Mol Cell Biol* 26 (15):5759-70.

Ting, C. N., M. C. Olson, K. P. Barton, and J. M. Leiden. 1996. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature* 384 (6608):474-8.

Ting, J. P., and J. Trowsdale. 2002. Genetic control of MHC class II expression. *Cell* 109 Suppl:S21-33.

Tolhuis, B., R. J. Palstra, E. Splinter, F. Grosveld, and W. de Laat. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10 (6):1453-65.

Townes, T. M., J. B. Lingrel, H. Y. Chen, R. L. Brinster, and R. D. Palmiter. 1985. Erythroid-specific expression of human beta-globin genes in transgenic mice. *Embo J* 4 (7):1715-23.

Triebel, F., and T. Hercend. 1989. Subpopulations of human peripheral T gamma delta lymphocytes. *Immunol Today* 10 (6):186-8.

Tuan, D., S. Kong, and K. Hu. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc Natl Acad Sci U S A* 89 (23):11219-23.

Vukmanovic, S., and F. R. Santori. 2005. Self-peptide/MHC and TCR antagonism: physiological role and therapeutic potential. *Cell Immunol* 233 (2):75-84.

Wai, A. W., N. Gillemans, S. Raguz-Bolognesi, S. Pruzina, G. Zafarana, D. Meijer, S. Philipsen, and F. Grosveld. 2003. HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. *Embo J* 22 (17):4489-500.

Wang, K., L. Gan, C. L. Kuo, and L. Hood. 1997. A highly conserved apoptotic suppressor gene is located near the chicken T-cell receptor alpha chain constant region. *Immunogenetics* 46 (5):376-82.

Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. et al., and E. S. Lander. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420 (6915):520-62.

West, A. G., M. Gaszner, and G. Felsenfeld. 2002. Insulators: many functions, many mechanisms. *Genes Dev* 16 (3):271-88.

West, A. G., S. Huang, M. Gaszner, M. D. Litt, and G. Felsenfeld. 2004. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell* 16 (3):453-63.

Wilson, A., W. Held, and H. R. MacDonald. 1994. Two waves of recombinase gene expression in developing thymocytes. *J Exp Med* 179 (4):1355-60.

Wilson, R. L., and G. V. Stauffer. 1994. DNA sequence and characterization of GcvA, a LysR family regulatory protein for the Escherichia coli glycine cleavage enzyme system. *J Bacteriol* 176 (10):2862-8.

Winoto, A., and D. Baltimore. 1989. A novel, inducible and T cell-specific enhancer located at the 3' end of the T cell receptor alpha locus. *Embo J* 8 (3):729-33.

Zhong, X. P., and M. S. Krangel. 1997. An enhancer-blocking element between alpha and delta gene segments within the human T cell receptor alpha/delta locus. *Proc Natl Acad Sci U S A* 94 (10):5219-24.

Zhong, X. P., and M. S. Krangel. 1999. Enhancer-blocking activity within the DNase I hypersensitive site 2 to 6 region between the TCR alpha and Dad1 genes. *J Immunol* 163 (1):295-300.