T Helper Cell Differentiation: Regulation by *cis* Elements and Epigenetics

Gap Ryol Lee,2,3,4 Sean T. Kim,2,3 Charalampos G. Spilianakis,2 Patrick E. Fields,2,5 and Richard A. Flavell1,2,*

1Howard Hughes Medical Institute
2Section of Immunobiology
Yale University School of Medicine
New Haven, Connecticut 06520

Cytokine loci undergo changes in chromatin structure when naive CD4+ T cells differentiate into Th1 or Th2 cells and have also been examined for regulatory sequences underlying such changes and their functional correlates. Studies have shown that distal regulatory elements control the *Ifng* and Th2 cytokine loci and are primary targets for tissue-specific transcription factors, serving as centers for epigenetic changes that mark heritable traits in effector cells. Reports of intra- and, remarkably, interchromosomal interactions between these regulatory elements shed light on the mechanisms by which they regulate gene expression, revealing an extraordinary new picture that conceptually extends our views on how genes are regulated from two to three dimensions. Here, we summarize these recent findings on the role of regulatory elements and their mechanisms of action, which are of broad significance for gene regulation, not only of the immune system but also of many, if not all, coregulated genes.

Introduction
Naive CD4+ T cells differentiate into Th1 or Th2 effector cells upon antigenic stimulation by antigen-presenting cells. Th1 and Th2 cells perform distinct immune functions; Th1 cells mediate cellular immunity against intracellular bacteria and viruses, whereas Th2 cells enable humoral immunity and immunity against extracellular parasites. The effector functions of Th1 cells are exerted in part by production of interferon (IFN)-γ and lymphotoxin-α and those of Th2 cells by interleukin (IL)-4, IL-5, and IL-13. Dysregulation of these Th1 and Th2 cell functions, however, can cause autoimmune disease and allergy, respectively.

The Th2 cytokine locus (*Il4-I13-Rad50-I15* locus) and *Ifng* locus, which are expressed in Th2 and Th1 cells, respectively, undergo structural and epigenetic changes during differentiation (Figures 1 and 2). These loci have been studied extensively as a model system for gene regulation and lineage-dependent chromatin changes because of fundamental interests in the differentiation process and their clinical implications. The studies include identification of regulatory sequences, analysis of their function, and elucidation of the mechanisms of gene regulation by these sequences. The Th2 cytokine locus, arguably the best-characterized system, presents a complex picture because of its sheer size (140 kb) and composition, containing diverse classes of regulatory elements (Figure 2).

Several excellent reviews have discussed transcription factors and lineage decision processes. In this review, we will focus on recent findings on the genetic analysis of key regulatory sequences in the Th2 cytokine and *Ifng* loci, their epigenetic modifications, and the mechanisms through which the regulatory elements regulate cytokine gene expression by intra- and interchromosomal interaction.

Regulatory Elements of the Th2 Cytokine and *Ifng* Loci
The search for regulatory elements involved in gene regulation has utilized three principal methods. In the first method, the DNase I hypersensitivity assay identifies genomic regions that have a relatively open chromatin structure due to a distorted nucleosomal array that confers susceptibility to digestion by DNase I. The second method involves the identification of conserved non-coding sequences (CNS) between different species. The rationale for this approach is that important regulatory sequences are likely to be conserved evolutionarily. The third approach is functional analysis. Commonly, the functional impact of elements discovered by DNase I hypersensitivity and sequence alignment is then determined in reporter assays, transgenic mice, and knock-out mice. In fact, functional analysis in bacterial artificial chromosome (BAC) transgenic mice was used to identify the Th2 locus control region (LCR), followed by DNase I hypersensitivity and sequence conservation analysis to localize and define the LCR elements.

Th2 Cytokine Locus
The search for regulatory elements in the Th2 cytokine locus using the methods described above has resulted in the identification of several different types of regulatory elements, including enhancers, silencers, and a locus control region. We will first describe each component briefly and discuss the overall roles of the elements in gene regulation.

Three hypersensitive sites (HS), termed HSS1–HSS3, were originally described in the intergenic region between *Il4* and *Il13* (Takemoto et al., 1998). Of the three, HSS1 and HSS2 are Th2 specific, whereas HSS3 is present in naive CD4+ cells, Th1, and Th2 cells. In a separate approach, Rubin and colleagues found highly conserved regions, including CNS-1 and CNS-2 (Loots et al., 2000); as it turned out, CNS-1 contained HSS1 and HSS2. To characterize the function of CNS-1, murine CNS-1 was deleted in the endogenous locus (Mohrs et al., 2001) and human CNS-1 was deleted in yeast artificial chromosome (YAC) DNA containing the Th2 cytokine locus (Loots et al., 2000). In both systems, deletion of CNS-1 resulted in a reduction of cells producing IL-4, IL-5, and IL-13, and it was suggested that CNS-1 is...
a coordinate regulator of Th2 cytokine genes. Subsequent studies in transgenic reporter mice showed that CNS-1/HSS has strong enhancer activity (Lee et al., 2001), whereas other experiments revealed that CNS-1 appeared to be responsive to GATA-3 in chromatin remodeling and enhancer activity, consistent with demonstration of potential GATA-3 sites by gel shift analysis (Takemoto et al., 2000; Lee et al., 2001).

HSV was identified as a Th2-specific HS located 3' of Il4 (Agarwal and Rao, 1998), overlapping with CNS-2. HVSa was described as an activation-dependent HS located close to HSV (Agarwal et al., 2000). In transient

Figure 1. Overview of Genetic, Epigenetic, and Molecular Events Accompanying T Helper Differentiation

General summary of epigenetic and structural changes of cytokine loci in naive, early activated, and effector Th1 and Th2 cells. Cytokine loci in naive CD4+ T cells exist in a relatively inert epigenetic state, with H3 hypoacetylation and DNA methylation at these sites. Both intra- and interchromosomal interactions between alternatively expressed loci are present. Upon antigenic stimulation, lineage nonspecific increases are seen at the epigenetic and transcriptional levels, with both IFN-γ and IL-4 mRNA being expressed. These changes are then reinforced and maintained in only the proper T helper subset, contributing to the polarization of effector Th1 and Th2 cells.

Figure 2. Detailed Map of Chromatin Structure throughout T Helper Differentiation

The Th2 cytokine and Ifng loci undergo lineage-specific chromatin changes during T helper differentiation. DNase I hypersensitive sites develop during lineage commitment, in part defining elements of regulatory function. Such hypersensitivity coincides with regions of histone acetylation and DNA demethylation, demarcating cis elements or areas of transcriptional activity. Depicted in the top panel is the entire 140 kb Th2 cytokine locus and the epigenetic changes that accompany T helper development. The bottom panel summarizes such changes occurring in the Ifng locus. Illustrated in this figure are hypersensitive sites and sites of histone H3 hyperacetylation and DNA demethylation as symbolized in the key. Genes are represented by thick, shaded bars with vertical lines protruding downward denoting the exons.
transfection and transgenic reporter assays, the combination of HSV and HSVa showed strong enhancer activity (Agarwal et al., 2000; Lee et al., 2001), where NFAT1 and GATA-3 bind to HSVa. HSV/CNS-2+HSVa knockout mice showed reduced expression of IL-4, IL-5, and IL-13 in Th2 cells and mast cells (Solymar et al., 2002), with the extent of reduction greater in mast cells. In Th2 cultures, there is a mixture of IL-4-producing and nonproducing cells, and interestingly, restriction enzyme accessibility of HSVa increases in the IL-4-producing cells, whereas it is unchanged in IL-4-nonproducers, suggesting that HSVa controls probabilistic expression of IL-4 (Guo et al., 2004).

The intronic enhancer (IE), located within the second intron of the If4, was initially described as a mast cell-specific HS (Henkel et al., 1992). It was later shown to be hypersensitive in Th2 cells, but not in Th1 or naïve CD4+ T cells (Agarwal and Rao, 1998). Transgenic reporter assays showed that IE/HSII enhanced IL-4 promoter activity and that IE in combination with CNS-1/HS3 conferred GATA-3-dependent enhancement of IL-4 promoter activity (Lee et al., 2001).

HS1/CS1/CGRE is located 1.6 kb upstream of Il13 and was described as a highly conserved Th2-specific HS (Agarwal and Rao, 1998; Kishikawa et al., 2001; Yamashtia et al., 2002). This region was shown to bind GATA-3, CBP/p300, and RNA polymerase II, and it enhanced Th2 cytokine promoter activity in transient transfection assays (Yamashtia et al., 2002). However, transgenic mice containing HS1-IL4 promoter luciferase did not show any enhancer activity compared to control IL-4 promoter-luciferase transgenic mice (G.R. Lee and R.A. Flavell, unpublished data), suggesting a discrepancy between the two assays.

HSIV is located at the 3’ end of If4 and is constitutively hypersensitive in naïve CD4+ T cells, Th1, and Th2 cells (Agarwal and Rao, 1998). Recently, the location of HSIV was finely mapped by using genomic DNA digested with restriction enzymes as molecular weight markers (Ansel et al., 2004). This newly revised location of HSIV (Ansel et al., 2004) is slightly different from that originally reported (Agarwal and Rao, 1998) and coincides with the location of HM1 (Lee et al., 2001). Unlike the originally assigned region of HSIV, the DNA sequence of HSIV/HM1 is highly conserved across many species (Ansel et al., 2004; Lee et al., 2001). Surprisingly, HM1 reduced IL-4 promoter activity, suggesting repressor/silencer activity (Lee et al., 2001). HSIV-deficient mice showed aberrant expression of IL-4 and IL-13 in Th1 cells, consistent with the function of HSIV as a Th1-specific silencer (Ansel et al., 2004).

An LCR is a regulatory region that comprises both enhancer and insulator activity; these confer high expression and insulation from the effects of neighboring chromatin on a given transgene. As a result, transgenic constructs carrying an LCR exhibit copy number-dependent tissue-specific expression. Because the Th2 locus comprises a cluster of cytokine genes (Il4, Il5, and Il13) that are regulated in a differentiation-specific manner, it was necessary to determine whether the Th2 cytokine locus contains an LCR. The presence of an LCR within the Th2 cytokine locus was first demonstrated with reporter assays using transgenic mice containing a BAC carrying the Th2 cytokine locus into which an IL-4 promoter-luciferase cassette was integrated (Lee et al., 2003). These transgenic mice showed copy number-dependent expression of luciferase as well as IL-4 and IL-13. Subsequent deletion analysis showed that the LCR was localized to the 3’ end of Rad50. DNase I hypersensitivity assays performed by our group and Rao and colleagues independently identified four HS sites (termed RHS4, RHS5/RAD50-O, RHS6/RAD50-(A+B), and RHS7/RAD50-C) that comprise the Th2 cytokine LCR (Fields et al., 2004; Lee and Rao, 2004). Among these sites, RHS4, 5, and 7 are Th2 specific, whereas RHS6 is constitutively expressed in naive CD4+ cells, Th1, and Th2 cells in our study. In contrast to our study, RHS5/RAD50-O is hypersensitive in both Th1 and Th2 cells in the study by Rao and colleagues; this difference is probably due to the use of primary cells in one study and cell lines in the other. RHS5, 6, and 7 are highly conserved between species, and functional analyses in transgenic reporter assays showed that the combination of these sites can recapitulate LCR function, suggesting that they constitute the core of the Th2 LCR (Fields et al., 2004). Among the HSs in the LCR, RHS7 was further analyzed in RHS7 knockout mice, which showed a dramatic reduction of Th2 cytokine expression, confirming the role of the LCR in Th2 cytokine gene expression (Lee et al., 2005).

In addition to meeting the classical definition of an LCR, the Th2 cytokine LCR exhibited other hallmarks that implicate a regulatory function for it. Most notably, the Th2 LCR undergoes a series of epigenetic changes in a lineage-specific pattern (Fields et al., 2004). The region encompassing RHS4-7 is hypoacetylated in naïve T cells. Upon TCR stimulation, this region becomes preferentially hyperacetylated under Th2-polarizing conditions, consistent with its increased hypersensitivity. Moreover, one specific site in the LCR, RHS7, undergoes a rapid and Th2-specific demethylation (Fields et al., 2004). Interestingly, these changes appeared to occur simultaneously with those observed at other important cis elements, in particular, the IL-4 and IL-13 promoters. These observations have yielded some of the first insights as to how coordinated regulation might be defined.

Ilng Locus

Three hypersensitive sites were described in the introns of the Ifng gene: HS1 and HS2 in intron 1 and HS3 in intron 3 (Agarwal and Rao, 1998). Transgenic reporter assays showed that a combination of intron 1 and intron 3 conferred strong enhancer activity, but not tissue-specific expression to an IFN-γ promoter reporter construct (Soutto et al., 2002). In addition, transgenic mice containing an 8.6 kb fragment spanning the human Ifng gene showed high level expression of the gene but lacked Th1 specificity, whereas a 191 kb fragment containing the human Ifng locus showed high level, Th1-specific expression, suggesting that distal regulatory elements are required for tissue-specific expression (Soutto et al., 2002; Young et al., 1989). Transient transfection assays showed enhancer activity of intron 3 (Shnyreva et al., 2004). HS1 and HS3 both become accessible at the chromatin level upon Th1 stimulation of CD4+ T cells as measured by increased restriction enzyme accessibility (Guo et al., 2004).
CNS-1 was discovered by the demonstration of DNA sequence conservation between species and is located 5 kb upstream of the *Ifng* gene (Lee et al., 2004; Shnyreva et al., 2004). CNS-1 demonstrated Th1-specific hypersensitivity and enhancer activity in transient transfection assays. NFAT and T-bet bind to this region in stimulated Th1 cell lines and augment enhancer activity (Lee et al., 2004). Use of the chromosome conformation capture (3C) method has recently shown that CNS-1 is colocalized with the *Ifng* gene in the naive and Th1 T cell nucleus (Spilianakis et al., 2005) because it shows significantly greater signal in the 3C assay than control GAPDH fragments only 560 bp apart (see below for a discussion of the 3C technique).

CNS-2 was also discovered by the analysis of DNA sequence conservation and is located 18 kb downstream of the *Ifng* gene (Shnyreva et al., 2004). CNS-2 showed enhancer activity in transient transfection assays, Th1-specific DNase I hypersensitivity, and Th1-specific histone acetylation. T-bet binds to CNS-2 and induces a transcriptionally favorable chromatin structure (Shnyreva et al., 2004). Again, 3C analysis has shown, interestingly, that CNS-2 becomes colocalized with the *Ifng* gene in Th1 cells, but these elements are not associated in precursor naïve T cells (Spilianakis et al., 2005).

**Epigenetic Control of Th1/Th2 Differentiation:**

**Chromatin Remodeling of the Cytokine Loci**

The regulatory sequences described above are the primary targets for binding tissue-specific transcription factors and chromatin remodeling factors and serve as platforms to initiate chromatin changes for transcriptional activation. We now will describe the basic principles of epigenetic changes and summarize the epigenetic changes that control Th1/Th2 differentiation (summarized in Figure 2).

The DNA/histone complex, called a nucleosome, forms the basic structural unit of chromatin. While this arrangement of chromatin (DNA wrapped around the cylinder-like histone core) enables the compaction of DNA, it simultaneously impedes recognition by DNA binding proteins. Thus, the structure is generally considered to be an impediment to transcriptional activity and must be “remodeled” in order to allow local access of transcriptional machinery. These changes accompany induction or silencing of transcription and include, but are not limited to, alterations in posttranslational modifications of histone tail residues and changes in DNA methylation status. Actively transcribed or transcriptionally competent gene loci generally exhibit increased sensitivity to nucleases or restriction enzymes, increased histone acetylation, and decreased CpG DNA methylation.

Although mechanisms of chromatin remodeling are multiple, an important contributing factor is the modification of histone tails, which protrude from the core nucleosomes and are accessible to a variety of modifying enzymes. Acetylation has been the most widely studied; however, recent work has correlated other modifications such as methylation, phosphorylation, and ubiquitination with changes in transcriptional activity in a number of genes (Cheung et al., 2000; Strahl and Allis, 2000; Wei et al., 1999; Baarends et al., 1999; El-Osta and Wolff, 2000; Turner, 2000). Those modifications of histone H3 that are positively correlated with transcriptional activity are acetylation of lysines 9 and 14 and methylation of lysine 4. The combination of these modifications is thought to demarcate “active chromatin domains” within a gene locus and is often associated with regulatory regions such as promoters and enhancers. In contrast, methylation of lysines 9 and 27 are associated with gene silencing.

The mechanisms by which histone modifications alter gene locus activity are not completely understood. One way in which this control is exerted is by altering nucleosomal architecture such that DNA binding is changed, either allowing or restricting access to DNA binding proteins. Another facet of this regulation may lie in the ability of a modified histone or CpG to serve as a template for binding of a variety of chromatin-modifying elements. This binding can occur in a combinatorial fashion and thus adds an additional layer of complexity to regulation by this mechanism (Strahl and Allis, 2000). This variety of permutations and potential regulatory scenarios is referred to as the “histone code.”

Another chromatin modification, CpG methylation of DNA, promotes gene silencing through a number of proposed mechanisms (Wilson et al., 2005). The methyl CpG binding proteins MBD2 and MeCP2 recruit complexes containing histone deacetylases (HDAC) and corepressors (Nan et al., 1998; Jones et al., 1998), and they may also prevent transcription factors and coactivators from binding to methylated promoters and enhancer regions. Thus, demethylation serves to derepress a locus, presumably through reversal of the aforementioned actions. Interestingly, HDAC recruitment by methyl-CpG binding proteins implicates an interplay between histone modifications and DNA methylation, perhaps alluding to the existence of a higher order system of combinatorial modifications beyond the histone code.

In early studies of DNA methylation of the *Ifng* gene, Young and colleagues made the first correlation between alterations in DNA structure and cytokine transcription (Farrar et al., 1985). In these studies, the prevention of DNA methylation by 5-azacytidine treatment led to expression of IFN-γ by cells that did not ordinarily express the cytokine. It was later demonstrated that changes in DNA methylation resulted not only in changes in transcription but also in alterations in the local chromatin environment surrounding that gene (Cedar 1988). Thirteen years after these studies, the first definitive evidence of chromatin changes playing a role in differential cytokine expression by T cell subsets was obtained (Agarwal and Rao 1998; Bird et al., 1998). Rao and colleagues identified T cell lineage-specific DNase I hypersensitive sites near the *Il4* and *Ifng* genes. These sites were found near promoter and enhancer regions in both gene loci (Agarwal and Rao 1998). Concurrent studies with HDAC inhibitors and 5-azacytidine identified a role for epigenetic events in regulation of transcription of the Th1/Th2 cytokines, confirming the importance of these events in T cell fate determination (Bird et al., 1998; Valapour et al., 2002).

That HDAC inhibition resulted in increased expression of IL-4 and IFN-γ provided indirect evidence for a role for histone acetylation in Th1/Th2 differentiation. More recent studies using chromatin immunoprecipitation (ChIP) directly demonstrated that changes in histone
acetylation at the cytokine loci accompany Th1/Th2 differentiation (Fields et al., 2002; Avni et al., 2002; Yamashita et al., 2002; Melli et al., 2003). While the cytokine loci in naive T cells are relatively unacetylated, hyperacetylation occurs on histones in the Th2 cytokine locus in Th2 cells and in the Il4g locus in Th1 cells after T cell stimulation. Time-course studies showed that initially the increased acetylation occurs on both loci independently of polarizing conditions. Lineage- and locus-specific patterns subsequently emerge and are reinforced by cytokines in a STAT-dependent manner. The hyperacetylation appears to be focused at promoters and other regulatory regions such as enhancers and the Th2 LCR. This suggested that the acetylation patterns “mark” the loci, maintaining their transcriptional competence in a lineage-specific fashion. This epigenetic mark persists in cells not actively transcribing the cytokines and thus may contribute to the maintenance of the cells’ identities as either Th1 or Th2. Studies in murine, memory CD4+ T cells support this hypothesis. Like effector T cells generated in vitro, the cytokine gene acetylation patterns in antigen-specific, effector, and central memory T cells showed Th1/Th2 polarization in vivo (Yamashita et al., 2002).

Although hyperacetylation correlates positively with transcriptional competence and may mark a gene locus for activity, its absence does not necessarily correlate with gene silencing. After polarizing freshly isolated, naive human T cells into Th1 or Th2, the cells were repolarized into the opposite phenotype. These repolarized cells showed increased acetylation at the formerly hyperacetylated cytokine locus, with no reduction of acetylation at the originally activated locus (Melli et al., 2003). These changes in acetylation correlated with plasticity in cytokine transcription. These results clearly demonstrated that hyperacetylation at an inactive cytokine gene locus can be reversed upon T cell stimulation under appropriate conditions and that this plasticity of locus activation enables the cell to be flexible in the type of responsiveness that it can mediate (Th1 versus Th2).

In a similar manner as hyperacetylation, DNA demethylation seems to correlate in general with the activation and transcriptional competence of cytokine gene loci as was found many years ago for other loci such as β-globin (van der Ploeg and Flavell, 1980). Various studies have established that in Th2 cells key cis elements in the Th2 cytokine locus undergo demethylation, including the Il4 and -13 promoters and CNS-1 and -2 (Fields et al., 2004; Lee et al., 2002; Guo et al., 2002; Santangelo et al., 2002). There are, however, areas of hypomethylation already present in naive T cells in the Il4 promoter (Fields et al., 2004) and CNS2 region (Lee et al., 2002). Once these sites of demethylation are established, they are maintained in a demethylated state and coincide with DNase hypersensitivity and hyperacetylation (Makar et al., 2003; Fields et al., 2002, 2004; Lee et al., 2002; Santangelo et al., 2002). Although it has not been shown that demethylation is sufficient to effect transcription, it is permissive (Busslinger et al., 1983), and ectopic expression of cytokine genes in 5-azacytidine-treated or MBD2-deficient T cells at least implicates a role for DNA methylation in maintaining a silenced state (Bird et al., 1998; Hutchins et al., 2002). MBD2-deficient Th1 and Th2 cells produce IL-4 and IFN-γ ectopically, and these conditions are exacerbated in the presence of the HDAC inhibitor trichostatin A (TSA), implicating a cooperative role of DNA methylation and histone hypoacetylation in silencing cytokine loci (Hutchins et al., 2002). Similarly, Dnmt1-deficient T cells exhibit increased ectopic production of cytokines in inappropriate cell types (Makar et al., 2003). Here, such derepression correlates with a general hypomethylation across the Il4 locus due to lack of Dnmt1 recruitment to the IL-4 promoter and intronic enhancer in Th1 cells. That GATA-3 expression is dispensable or unaffected in these knockout animals demonstrates an elegant mechanism of bypassing conventional transcriptional programs to activate cytokine loci solely through epigenetic alterations.

There is increasing evidence for the role of DNA methylation in governing cytokine expression developmentally before and after T helper differentiation. For example, a pattern of hypomethylation at the IFN-γ promoter is established early in T cell development as well as in naive CD8+ and NK populations, but not in B or nonhematopoietic cells. This hypomethylation, indicative of a predisposition for IFN-γ production, is maintained in Th1 cells, but this region becomes remethylated during Th2 development, conditions under which IFN-γ is not expressed (Yano et al., 2003; Winders et al., 2004). Here, methylation of DNA silences a locus as a function of terminal differentiation. On the other hand, the newly identified CIRE (conserved intron 1 sequence of Il4) undergoes demethylation only in IL-4-secreting Th2 cells, which the authors argue explains in part the basis for memory expression of IL-4 (Tykocinski et al., 2005). Thus, DNA demethylation also potentiates heritable and stable gene expression during the transition from effector to memory T cell.

Very little is known about the transcription factors that mediate epigenetic changes at the cytokine loci. In the Il4 and Ifng loci, cytokine signaling via the STAT proteins has been shown to be a critical determinant of polarized acetylation (Avni et al., 2002; Fields et al., 2002). In the absence of STAT6, neither Il4 locus histone acetylation nor DNA demethylation is completely established (Fields et al., 2002; Lee et al., 2002). Likewise, in STAT4-deficient T cells, IFN-γ promoter acetylation does not occur. Although these proteins have been shown to bind directly to the gene loci, whether their effects result from recruitment of histone acetyltransferase or from indirect action has not yet been determined (Morinobu et al., 2004; Avni et al., 2002). The roles of the downstream activators GATA-3 and T-bet appear to be more complicated than that of the STATs. Enforced, ectopic expression of these proteins is sufficient to induce chromatin changes in the cytokine loci. For example, GATA-3 can induce DNase I hypersensitivity and targeted hyperacetylation in the Il4 locus (Ouyang et al., 2000; Takemoto et al., 2000; Lee et al., 2000; Fields et al., 2002). Likewise, overexpression of T-bet can induce similar changes in the Ifng locus (Fields et al., 2002, Mullen et al., 2001). Whether these act directly on the loci or indirectly has not been established. The necessity of these factors in inducing acetylation changes has been examined by using T cells with targeted mutations of these factors. In T-bet-deficient T cells, which are defective in developing Th1 responses,
hyperacetylation in the IFN-γ promoter fails to develop (Avni et al., 2002). In conditional GATA-3-deficient T cells, normal Th2 acetylation patterns in the Il4, Il5, and Il13 genes are impaired (Yamashita et al., 2004). The effect is most prominent in the Il5 gene, where there is a severe loss of Th2-associated hyperacetylation. The effect in the Il4 and Il13 genes is much less pronounced.

The effect of GATA-3 loss on Th2 cytokine production is more apparent in developing responses than in established Th2 cells. In a developing Th2 response, loss of GATA-3 results in the inability to effectively mount a Th2 response, with a substantial reduction in the generation of IL-4-, IL-5-, and IL-13-producing cells (Pai et al., 2004; Zhu et al., 2004; Yamashita et al., 2004). In contrast, the loss of GATA-3 in established Th2 cells causes a dramatic reduction in IL-5 production, a smaller effect on IL-13, but little effect on IL-4 (Pai et al., 2004; Zhu et al., 2004). GATA-3 acts in part by maintaining hyperacetylation of promoter regions of the Il5 gene. The maintenance of acetylation patterns and transcriptional activity of the Il4 and Il13 genes may be controlled by additional mechanisms.

While acetylation has been extensively studied, other histone modifications in the cytokine loci have recently begun to be analyzed. In particular, histone H3 methyltransferase Ezh2 with the same region. Thus, correlated with this was the localization of the H3/K27 methylation at the molecular level. Technological advances in determining the spatial organization of chromatin have enabled the study of the dynamic changes that occur during cell differentiation. Recent studies of the Th2 cytokine and Ilfng loci using this technology have highlighted the role of the Th2 LCR and shed light on its mechanism of locus regulation.

**LCRs and Their Mode of Action**

LCRs have three major characteristics as defined in transgenic reporter assays. They confer tissue-specific expression of linked genes at physiological levels in a position-independent, copy number-dependent manner. LCRs have the ability to incorporate the action of an enhancer as well as an insulator and that is how they are able to activate a gene irrespective of the site of integration in the genome. The components of an LCR usually colocalize with a series of DNase I hypersensitive sites in the chromatin of expressed genes. These sites are occupied by ubiquitously expressed as well as tissue-specific factors.

Numerous DNA elements have been described in mammals that meet the criteria for LCR function (Li et al., 2002). The concept, discovery, and definition of the LCR derives from pioneering studies in the β-globin locus (Grosveld et al., 1987) and still most of the data presented so far for LCR function come from studies of the β-globin LCR (human, murine, and chicken) (Festenstein and Kioussis, 2000; Fraser and Grosveld, 1998, 1999; Li et al., 1999b). This LCR consists of five DNase I hypersensitive sites, the first four of which are erythroid specific and the fifth is found in multiple cell lineages (Li et al., 1999a). The first application of 3C analysis to mammalian genes (see below) identified an “active chromatin hub” at which the LCR hypersensitive sites contact the promoter of the gene being expressed at a given time via loops. For the β-globin locus, this occurs in a developmentally regulated manner, such that at later times a new gene/promoter is recruited to the hub, presumably enabling this gene to be expressed (Palstra et al., 2003; Tolhuis et al., 2002). Other examples include those of the human cd2 locus (Festenstein et al., 1996; Lang et al., 1991), the T cell-specific TCRα LCR (Diaz et al., 1994), which confers tissue-specific DNA demethylation, the murine immunoglobulin heavy chain (Igh) LCR (Madisen and Groudine, 1994), and many others.

**The Th2 LCR and the Coordinate Expression of Three Genes**

The discovery of the β-globin LCR and the characterization of LCRs in other systems like that of the Th2 cytokine locus suggest that tissue-specific, developmentally regulated gene transcription relies not only on gene-proximal elements but also on long-range interactions of various cis-regulatory elements and complex secondary chromatin configurations (Li et al., 2002). Although defined by many criteria, there was no clear mechanism suggested for the mode of action of the Th2 LCR in coordinately regulating the three Th2 cytokine genes.

Again derived from the paradigm of the globin model, according to the looping model (Choi and Engel, 1988), an LCR is occupied by activating transcription factor complexes that come in close proximity to a promoter and activate a gene by looping out the intervening chromosomal region between the LCR and the promoter. The first elucidation of the active chromosome hub for the β-globin locus by 3C analysis provided a molecular basis for LCR function in regulation of gene expression during development (Palstra et al., 2003; Tolhuis et al., 2002). Thus, LCRs can confer their function even when located at great distances from their target gene(s) (Dekker, 2003). However, how the expression of three coregulated genes—Il4, Il5, and Il13—was controlled by an LCR was unclear. Using 3C (Tolhuis et al., 2002), it was shown that the Th2 cytokine locus forms an initial core complex in which the Th2 LCR can exert its activating effects by direct physical interaction (Spilianakis and Flavell, 2004). In naive T cells and NK cells (which share a common lymphoid progenitor), a cell-specific chromatin core is formed between the participating Th2 genes (Il4, Il5, and Il13), the enhancers in the locus (including CNS1), and the Th2 LCR; in contrast, this conformation does not exist in B cells or fibroblasts. Earlier studies on LCR function showed enhancement of IL-4 and IL-13, but not IL-5 (Lee et al., 2003). This result was confirmed by the strong interaction of the LCR with the Il4

**The Dynamic Changes That Occur during Cell Differentiation.** Recent studies of the Th2 cytokine locus (Festenstein et al., 1999; Li et al., 1999a) have provided a molecular basis for LCR function in regulation of gene expression during development (Palstra et al., 2003; Tolhuis et al., 2002). Thus, LCRs can confer their function even when located at great distances from their target gene(s) (Dekker, 2003). However, how the expression of three coregulated genes—Il4, Il5, and Il13—was controlled by an LCR was unclear. Using 3C (Tolhuis et al., 2002), it was shown that the Th2 cytokine locus forms an initial core complex in which the Th2 LCR can exert its activating effects by direct physical interaction (Spilianakis and Flavell, 2004). In naive T cells and NK cells (which share a common lymphoid progenitor), a cell-specific chromatin core is formed between the participating Th2 genes (Il4, Il5, and Il13), the enhancers in the locus (including CNS1), and the Th2 LCR; in contrast, this conformation does not exist in B cells or fibroblasts. Earlier studies on LCR function showed enhancement of IL-4 and IL-13, but not IL-5 (Lee et al., 2003). This result was confirmed by the strong interaction of the LCR with the Il4...
and \( \text{II13} \) genes, but not with the \( \text{II5} \) gene as shown by 3C analysis (Spilianakis and Flavell, 2004). Upon differentiation, Th2 cells express cell-specific transcription factors that are rapidly recruited to the preformed chromatin complex generated by the Th2 LCR and the genes it regulates. Thus, this provides an explanation in molecular terms as to how an LCR can coordinately regulate three genes in a given cell type. The Th2 LCR can, in this way, coactivate the three cytokine genes participating in a preformed chromatin configuration with- out affecting the constitutively expressed \( \text{Rad50} \) gene, which is looped out of the active chromatin hub. It remains to be seen whether cell-specific transcription factors such as GATA-3 and STAT6 act as the platform for the recruitment of remodeling complexes or other proteins and thus confer specificity of recruitment of these complexes to a locus where transcription is ready to be initiated.

Prepoised and Poised Gene Chromatin Configuration in the Th2 Cytokine Locus

These studies have shown that the promoters of the Th2 cytokine genes come into close spatial proximity to form a “prepoised” core chromatin configuration in cells of the T cell lineage that either express (Th2) or do not express (naive T cells, Th1) these cytokine genes (Spilianakis and Flavell, 2004). This configuration is also found in NK, B cells, and fibroblasts. Within this core complex in cells of the T and NK cell lineages, the LCR further contributes to this higher order chromatin conformation to generate a “poised” state. It has already been described for several systems that a transcriptional complex can be formed on a promoter, but transcription is not initiated because the C-terminal domain of RNA polymerase II is not phosphorylated (Boehm et al., 2003; Gilmour and Lis, 1986; Graunke et al., 1999; Ljungman and Hanawalt, 1995; Spilianakis et al., 2003; Wang et al., 1997). The \( \beta \)-globin LCR acquires such a conformation, interacting with other DNase I hypersensitive sites in the locus in erythroid progenitors that do not express the \( \beta \)-globin genes (Palstra et al., 2003). The phenomenon in which a locus is heritably poised for gene expression even when it is transcriptionally inactive may be a unique feature of genes that require rapid transcriptional induction in response to a stimulus (Smale and Fisher, 2002). CD4+ naive T cells, Th1, and Th2 cells initially have the potential to express Th2 cytokines with regard to their chromatin conformation, but additional cell-specific factors and posttranslational histone modifications of chromatin determine whether a given event occurs.

Transcription Factors Important for the Action of the Th2 LCR

Naive CD4+ T cells produce low levels of IL-4 in a STAT6-independent manner, (Grogan et al., 2001) with peak mRNA production occurring 2–3 hr poststimulation.
We hypothesize, taking into consideration recent data regarding the looped conformation of the Th2 cytokines, that early expression of the Th2 cytokines is supported by the initial poised chromatin configuration, but in later developmental stages, such as in polarized effector cells, STAT6 is important for the sustained tissue-specific expression of the Th2 cytokines, where additional factors may be needed to mark and maintain the memory of epigenetic modifications. Thus, STAT6 is necessary for the formation or maintenance of a secondary chromatin conformation in later stages of differentiation, but other factors are required to generate the prepoised conformation already found in nondifferentiated naive cells.

Another downstream effector of STAT6 is GATA-3, a factor important for the expression of Th2 cytokines (Zheng and Flavell, 1997) and the formation and/or maintenance of interactions between the Th2 LCR and the Th2 cytokine gene promoters (Lee et al., 2005). We suggested that GATA-3 might coordinate affect the expression of Il4, Il5, and Il13 in the Th2 cytokine locus, exerting its action on an already preformed, prepoised, chromatin configuration (Spilianakis and Flavell, 2004). Moreover, because GATA-3 can induce chromatin remodeling activity (Ouyang et al., 2000; Takemoto et al., 2000, 2002; Lee et al., 2000), it may facilitate the formation or maintenance of the interaction between the Th2 LCR and the promoters of the target cytokine genes. GATA-3 may cooperate with other tissue-specific transcription factors for these functions. One example is cooperative binding with NFAT by TCR stimulation. TCR stimulation followed by calcium mobilization leads to the rapid dephosphorylation of NFAT proteins and their subsequent translocation to the nucleus where they exhibit increased affinity for DNA binding (Rao et al., 1997). Transcription-activating protein complexes consist of NFAT dimers and NFAT/GATA family members heterodimers among other components (Okamura and Rao, 2001).

Interchromosomal Interaction of the Th2 and Ifng Loci

Regulation of gene loci from genetic regulatory elements is based on the binding of transcription factors and chromatin remodeling proteins that may bring into close proximity genes whose expression has to be coordinately regulated.

As discussed previously, upon stimulation of naive CD4+ T cells and prior to their final differentiation into Th1 or Th2 cells, an early burst of transcription occurs for both the Il4 and Ifng genes that is dependent on stimulation but independent of cytokine signaling (Ansel et al., 2004; Grogan et al., 2001; Guo et al., 2004; Spilianakis et al., 2005). This observation led us to utilize the 3C technique to test the hypothesis that this could be mediated via physical interactions between the Ifng locus and the Th2 LCR. We found that the Ifng gene, among other interactions, physically interacts with RHS6 of the Th2 LCR. These data were also confirmed at the single cell level by using fluorescent in situ hybridization (FISH), showing that in almost 40% of all naive CD4+ T cells, the two loci were seen in close proximity (Spilianakis et al., 2005). Upon differentiation of helper T cells under either Th1 or Th2 conditions, the interactions between the loci wane and interchromosomal interactions are favored. Resting naive CD4+ T cells do not express IFN-γ or Th2 cytokines, but upon activation, they express both rapidly without the presence of key transcription factors. We therefore speculate that the interactions between these two loci most likely hold the Ifng and Il4 gene complexes together in a poised state, ready for transcription, but not allowing active transcription to take place.

It was recently reported that GATA-3 and T-bet are expressed at low but detectable protein levels in naive CD4+ cells and physically interact through tyrosine phosphorylation of T-bet as shown in communoprecipitation assays (Hwang et al., 2005). This study was yet another example of an interaction, at the protein level, of two products with exclusive function in Th1 and Th2 cells. It would be interesting to investigate the possibility that these two factors are coupled in naive T cells and bind at the Th2 (GATA-3) and Th1 (T-bet) loci to provide the interaction that holds these two chromosomes together. This could be tested in mice lacking either GATA-3 and/or T-bet and provide a clear answer for the involvement of these two key transcription factors in this phenomenon. Of major interest would also be the characterization of the protein complexes bound to the Th2 LCR and the identification of protein complexes with specific remodeling activity acting at long distances in chromatin rather than at promoter and enhancer elements. Among many approaches, yeast one-hybrid assays utilizing DNA baits that span the different DNase I hypersensitive sites of the Th2 LCR could be used to identify such protein complexes.

Future Directions

It remains to be seen how the inter- and intrachromosomal cytokine gene complexes that we have described converge with epigenetic changes to effect gene transcription. We know that many of these events happen concurrently, but it is unknown whether paradigms observed in other LCRs are mirrored in the cytokine gene system. For example, deletion of a specific HS in the TCRα LCR precludes demethylation of a downstream HS (Santoso et al., 2000). Do certain elements of the Th2 LCR enable epigenetic changes in other regulatory elements in the locus? Conversely, do any of these changes allow for the association or dissociation of genetic loci?

As discussed in this review, several approaches have been employed to understand gene regulation in the immune system and specifically the mechanisms of Th1/Th2 differentiation. First was the characterization and genetic manipulation of cis-acting elements and trans-acting transcription factors that regulate T helper cell expression patterns. Second, recent advances in dissecting the histone code and the regulation of DNA methylation have added one more level of complexity but have also helped to elucidate complex gene activation pathways. Third, the recent elucidation of the organization of these loci in the nucleus and the surprising and unprecedented discovery of interactions between regulatory components of cytokine loci on different chromosomes have bred exciting new areas for study of this interesting problem. The implications of the finding of interchromosomal interactions are significant. We
can now consider regulatory elements like LCRs, enhancers, and silencers not only as regulatory elements with special features acting in cis to drive the expression of certain genes but also as genetic elements that can act across multiple chromosomes in trans with pleiotropic functions such as nuclear relocalization of loci and the regulation of gene expression not only on the same chromosome but also on different chromosomes. Many exciting issues remain to be resolved: are these interactions stable or transient, and are there alternative configurations for a given element like the Th2 LCR? The next years will yield these answers. It is highly likely that this kind of interaction will not just be a property of cytokine genes but, instead, a general feature of gene organization. Interesting candidates for interchromosomal gene regulation are the immunoglobulin heavy and light chain genes, where rearrangement must be controlled in a temporal and locus-specific fashion, the \( \alpha- \) and \( \beta- \) globin genes, where the level of expression of two genes must be balanced and the genes expressed matched to the developmental stage, and olfactory receptor genes, where a choice to express a single gene among many located on multiple chromosomes must be made. Finally, a potential contribution of these kinds of interactions between different chromosomes to the common chromosomal translocations seen in tumors should be explored.

Acknowledgments

We would like to thank E. Eynon for critical review of this manuscript and F. Manzo and R. Champion for assistance with manuscript preparation. Part of this work was supported by a grant from the National Institutes of Health (NIH) (HL56389). C.G.S. is supported with a Can-

References


mn GADD45 gene for rapid transcriptional up-regulation. Nucleic Acids Res. 27, 3881–3890.


Guo, L., Hu-Li, J., Zhu, J., Watson, C.J., Difilipantionio, M.J., Pannet-


Hutchins, A.S., Mullen, A.C., Lee, H.W., Sykes, K.J., High, F.A., Hen-


