Towards an understanding of cell-specific functions of signal-dependent transcription factors

Dawn X Zhang1,2 and Christopher K Glass1,3

1Department of Cellular and Molecular Medicine 2Biomedical Sciences Graduate Program 3Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

Abstract

The ability to regulate gene expression in a cell-specific manner is a feature of many broadly expressed signal-dependent transcription factors (SDTFs), including nuclear hormone receptors and transcription factors that are activated by cell surface receptors for extracellular signals. As the most plastic cells of the hematopoietic system, macrophages are responsive to a wide spectrum of regulatory molecules and provide a robust model system for investigation of the basis for cell-specific transcriptional responses at a genome-wide level. Here, focusing on recent studies in macrophages, we review the evidence suggesting a model in which cell-specific actions of SDTFs are the consequence of priming functions of lineage determining transcription factors. We also discuss recent findings relating lineage-determining and SDTF activity to alterations in the epigenetic landscape as well as the production and function of enhancer RNAs. These findings have implications for the understanding of how natural genetic variation impacts cell-specific programs of gene expression and suggest new approaches for altering gene expression in vivo.

Key Words

- cell-specific
- lineage determining transcription factors
- signal dependent transcription factors
- genome-wide analyses

Introduction

A central question in molecular endocrinology is how a single hormone-responsive transcription factor, reading the same DNA template, is capable of regulating different genes in different cell types. For example, activation of the glucocorticoid receptor by endogenous or synthetic glucocorticoids promotes a gluconeogenic program of gene expression in hepatocytes; mediates negative feedback control of adrenocorticotropic hormone secretion from the corticotrophs of the anterior pituitary; and suppresses pro-inflammatory responses of macrophages (Lazar 2011). In each case, the same glucocorticoid receptor controls the expression of a distinct set of genes that are crucial to the specific biological functions of the corresponding cell type. This ability to regulate gene expression in a cell-specific manner is shared by many broadly expressed signal-dependent transcription factors (SDTFs), including other nuclear hormone receptors and members of the diverse families of transcription factors that are activated by cell surface receptors for extracellular signals (e.g., STAT transcription factors, NF-κB family members, CREB, etc.). (For a complete listing of transcription factor names and abbreviations, see Table 1). The development of high-throughput sequencing methods to characterize transcription factor binding, histone modifications, and nascent RNA production at the genome-wide level has provided a powerful array of new tools for investigating the molecular basis for cell-specific transcriptional responses. These tools have been applied in a
number of biologically important model systems, including models for cellular differentiation (e.g., embryonic stem cell differentiation, adipogenesis, and hematopoiesis) and signal-dependent gene expression (e.g., macrophage activation and hormone-dependent breast and/or prostate cancers). In this review, we focus on recent insights into the molecular basis of cell-specific functions of broadly expressed SDTFs derived from genome-wide studies on macrophages.

Macrophages were discovered in 1884 by Ilya Mechnikov, a Russian bacteriologist who later shared the 1908 Nobel Prize in Physiology or Medicine with Paul Ehrlich for their studies in innate immunity (Buryachkovskaya et al. 2013). Classically, macrophages differentiate from bone marrow-derived monocytes; however, recent studies have found that at least some tissue-resident macrophages such as microglia, Langerhans cells, and Kupffer cells are derived from hematopoietic stem cells during fetal development (Saijo & Glass 2011, Wynn et al. 2013). As mediators of innate immunity and normal homeostatic processes, macrophages are essential to the body’s ability to control inflammation. Consequently, these immune cells have been implicated in multiple disease processes such as diabetes, rheumatoid arthritis,

### Table 1
Abbreviations and full names of transcription factors cited in this review

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Transcription factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU.1 or SPI1</td>
<td>Spleen focus-forming virus (SFFV) proviral integration oncogene SPI1</td>
</tr>
<tr>
<td>AP1</td>
<td>Activator protein 1 (heterodimeric protein complex)</td>
</tr>
<tr>
<td>C/EBPs (CEBPα and CEBPB)</td>
<td>CCAAT-enhancer-binding proteins (α and β)</td>
</tr>
<tr>
<td>SOX2</td>
<td>SRY (sex-determining region Y)-box 2</td>
</tr>
<tr>
<td>FOXD3</td>
<td>Forkhead box D3</td>
</tr>
<tr>
<td>JUN</td>
<td>Jun proto-oncogene</td>
</tr>
<tr>
<td>FOS</td>
<td>FBJ murine osteosarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>TCF3</td>
<td>Transcription factor 3</td>
</tr>
<tr>
<td>EBF1</td>
<td>Early B-cell factor 1</td>
</tr>
<tr>
<td>FOXO1</td>
<td>Forkhead box O1</td>
</tr>
<tr>
<td>LXR</td>
<td>Liver X receptor</td>
</tr>
<tr>
<td>NR1D1 and NR1D2</td>
<td>Nuclear receptor subfamily 1, group D, members 1 and 2 (Rev-Erbβ and β)</td>
</tr>
<tr>
<td>RELA</td>
<td>V-rel reticuloendotheliosis viral oncogene homolog A (avian)</td>
</tr>
<tr>
<td>SRF</td>
<td>Serum response factor</td>
</tr>
<tr>
<td>STAT5 (STAT1, STAT3, etc.)</td>
<td>Signal transducer and activator of transcription (1, 3, etc.)</td>
</tr>
<tr>
<td>SMAD3</td>
<td>Mothers against decapentaplegic homolog 3 or SMAD family member 3</td>
</tr>
<tr>
<td>HNF4A</td>
<td>Hepatocyte nuclear factor 4, α</td>
</tr>
<tr>
<td>FOXA1</td>
<td>Forkhead box A1</td>
</tr>
<tr>
<td>JUNB</td>
<td>Jun B proto-oncogene</td>
</tr>
<tr>
<td>IRF4</td>
<td>Interferon regulatory factor 4</td>
</tr>
<tr>
<td>ATF3</td>
<td>Activating transcription factor 3</td>
</tr>
<tr>
<td>RUNX1</td>
<td>Runt-related transcription factor</td>
</tr>
<tr>
<td>CBFβ</td>
<td>Core-binding factor, β subunit</td>
</tr>
<tr>
<td>ETS1</td>
<td>V-ets erythroleukemia virus E26 oncogene homolog 1 (avian)</td>
</tr>
<tr>
<td>ELF1</td>
<td>E74-like factor 1 (ets domain transcription factor)</td>
</tr>
<tr>
<td>FOXP3</td>
<td>Forkhead box P3</td>
</tr>
<tr>
<td>NFAT</td>
<td>Nuclear factor of activated T-cells</td>
</tr>
<tr>
<td>BATF</td>
<td>Basic leucine zipper transcription factor, ATF-like</td>
</tr>
<tr>
<td>RORG</td>
<td>RAR-related orphan receptor C</td>
</tr>
<tr>
<td>MAF</td>
<td>V-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)</td>
</tr>
<tr>
<td>EP300</td>
<td>E1A binding protein p300</td>
</tr>
<tr>
<td>GATA3 and GATA4</td>
<td>GATA-binding protein 3 and 4</td>
</tr>
<tr>
<td>TBX21</td>
<td>T-box 21</td>
</tr>
<tr>
<td>POU5F1 (OCT4)</td>
<td>POU class 5 homeobox 1</td>
</tr>
<tr>
<td>MYOD1</td>
<td>Myogenic differentiation 1</td>
</tr>
<tr>
<td>ESR1</td>
<td>Estrogen receptor 1</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>TFAP2C</td>
<td>Transcription factor AP2γ (activating enhancer binding protein 2γ)</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid receptor</td>
</tr>
<tr>
<td>MLL (1/3, 2/4, 3, etc.)</td>
<td>Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, <em>Drosophila</em>; 1/3, 2/4, 3, etc.)</td>
</tr>
<tr>
<td>LSD1</td>
<td>Lysine (K)-specific demethylase 1A</td>
</tr>
</tbody>
</table>

http://jme.endocrinology-journals.org

DOI: 10.1530/JME-13-0216

© 2013 Society for Endocrinology

Printed in Great Britain

Published by Bioscientifica Ltd

Downloaded from Bioscientifica.com at 11/17/2018 12:26:15AM
via free access
multiple sclerosis, and atherosclerosis (Drexler & Foxwell 2010, Glass & Saijo 2010, Olefsky & Glass 2010). Primary macrophages are easily obtainable for in vitro studies (as bone marrow-derived macrophages, thioglycollate-elicited macrophages, circulating monocytes, splenic monocytes, etc.) from wild type and genetically modified animals and can also be differentiated from human blood monocytes. The ability to obtain large numbers of these cells makes them a robust model system for contemporary ‘omics’ technologies – proteomics (Hettinger et al. 2013, Meissner et al. 2013), lipidomics (Spann et al. 2012, Maurya et al. 2013), and genomics (Heinz et al. 2010, Kaikkonen et al. 2013, Lam et al. 2013, Ostuni et al. 2013) – for investigating signal-dependent mechanisms.

As the most plastic cells of the hematopoietic system, macrophages are responsive to many regulatory molecules, including growth factors, pro- and anti-inflammatory cytokines, pathogen-associated molecular patterns (PAMPs), damage-associated molecular patterns (DAMPs), and nuclear receptor ligands (Smale 2010, Lawrence & Natoli 2011, Murray & Wynn 2011, Shi & Pamer 2011, Sica & Mantovani 2012, Van Dyken & Locksley 2013, Wynn et al. 2013). Responses to PAMPs and DAMPs are mediated by pattern recognition receptors that include the family of toll-like receptors (TLRs; Medzhitov et al. 2011). TLR4 recognizes lipopolysaccharide (LPS), a component of the cell wall of Gram-negative bacteria. Upon binding LPS, TLR4 signal transduction activates a number of latent transcription factors, which include NF-kB and interferon regulatory factors (Medzhitov et al. 1997, Poltorak et al. 1998a,b). These factors subsequently enter the nucleus and regulate the expression of inflammatory response genes, which play essential roles in innate immunity and contribute to the development of adaptive immunity (Medzhitov & Horng 2009). This transcriptional response is among the most dramatic in biology with respect to both the number of genes regulated (thousands) and the magnitude of change in gene expression (> 4 orders of magnitude for the most highly regulated genes) (Escoubet-Lozach et al. 2011, Kaikkonen et al. 2013). This response is also subject to counter-regulation by anti-inflammatory signaling molecules that include ligands for nuclear hormone receptors, such as the glucocorticoid receptor, peroxisome proliferator-activated receptors (PPARs), and liver X receptors (LXRs) (Glass & Saijo 2010). These properties of the macrophage make it a particularly rich cell type for the application of genome-wide approaches.

Recent studies of signal-dependent transcriptional responses in macrophages and B cells have suggested a
general model to explain many of the cell-specific functions of broadly expressed SDTFs (Heinz et al. 2010; Fig. 1). In this model, relatively simple combinations of lineage-determining transcription factors (LDTFs) are proposed to function collaboratively to select genomic enhancer-like regulatory regions in a cell-specific manner. The collaborative binding of these factors results in nucleosome remodeling to generate open regions of chromatin that provide access to SDTFs. Thus, SDTFs are directed to the vicinity of target genes in a cell-specific manner as a consequence of the priming functions of the lineage-determining factors for that particular cell type.

**Setting the stage: roles of lineage determining factors**

It has long been proposed that the functional output of transcription factors is determined by their competitive binding with nucleosomes. Until nucleosomes are displaced and chromatin regions are open, most transcription factors bind transiently to chromatin in a manner insufficient for efficient transcription (Lickwar et al. 2012). For example, during hematopoietic stem cell differentiation, nucleosome remodeling by the ATP-dependent, switch/sucrose non-fermentable (SWI/SNF) chromatin-remodeling complex allows for transcription factor binding at erythroid LDTF binding sites (Hu et al. 2011). Furthermore, studies utilizing novel high-throughput sequencing methods have been used to define open chromatin regions, which contain important transcriptional regulatory elements that delineate cell-specific identity (Song et al. 2011).

Recent studies of the ENCODE consortium estimate that the mammalian genome contains hundreds of thousands of enhancers (Bernstein et al. 2012), the majority of which are selected for activity in a cell-specific manner. This process of enhancer selection is proposed to result from collaborative interactions of LDTFs, also referred to as pioneer factors or master regulators. These factors are sequence-specific DNA-binding proteins with the ability to access their binding sites even when those sites are wrapped in nucleosomes. The factors also represent placeholders that can be replaced by other transcription factors at later stages during development (Zaret & Carroll 2011). For example, in embryonic stem cells, developmental pioneer transcription factors, SOX2 and FOXD3, bind at tissue-specific elements, and are replaced by cell-specific transcription factors after differentiation (Liber et al. 2010). Currently, two prevailing theories exist to explain how pioneer transcription factors function in defining cell-specificity during development. In the passive model, pioneer transcription factor occupancy decreases the number of binding events necessary for transcriptional activation (Zaret & Carroll 2011). In contrast, the active model suggests that pioneer transcription factors function by recruiting chromatin-remodeling complexes to activate transcription (Zaret & Carroll 2011).

The roles of LDTFs in priming cell-specific regulatory sites have been well-characterized in cells of the hematopoietic lineage. The LDTF, spleen focus-forming virus (SFFV) proviral integration oncogene, SPI1, more commonly referred to as PU.1, provides a particularly instructive example of an LDTF that is required for specification of more than one hematopoietic cell type; genetic deletion of PU.1 results in loss of macrophages, neutrophils, and B cells (Scott et al. 1994, 1997). In myeloid lineage precursor cells, PU.1 instructs hematopoietic progenitors to upregulate myeloid-specific cell surface antigens and to downregulate other cell-specific markers and transcription factors (Nerlov & Graf 1998). Recent studies have shown a dependence on cell cycle lengthening and subsequent PU.1 accumulation, which dictates myeloid differentiation from the common myeloid lymphoid progenitor (Kueh et al. 2013). In lymphoid progenitor cells, PU.1 is required for the expression of an alternative set of genes required for the progression to mature B cells (Lin et al. 2010). Thus, even a single LDTF can promote distinct programs of gene expression in different cell types.

PU.1 is a member of the E-twenty six (ETS) family of transcription factors, which bind to the canonical ETS-motif, 5′-GGAA-3′. Using genetic and genomic methods, PU.1 was recently shown to select macrophage or B cell-specific enhancers based on the co-occurrence of nearby binding sites for either macrophage or B-cell LDTFs, which are selectively expressed in one or the other cell type (Heinz et al. 2010). In macrophages, PU.1 binding occurred at enhancer-like regions exhibiting nearby binding sites for other essential myeloid LDTFs, CCAAT/enhancer binding protein κ and β (C/EBPs), and/or API transcription factors, a heterodimeric protein composed of Jun proto-oncogene (JUN) and FBj murine osteosarcoma viral oncogene homolog (FOS) (Heinz et al. 2010). In B cells, PU.1 binds at enhancer-like regions containing binding sites for other B cell LDTFs including TCF3, EBF1, and FOXO1 (Heinz et al. 2010, Lin et al. 2010). Notably, binding of PU.1 and alternative LDTFs was mutually dependent; genetic deletion of one factor resulted in loss of binding of the other at closely spaced sites (Heinz et al. 2010). Although generally within 100 bp
of each other, no strict spacing relationship was observed between PU.1 and collaborative transcription factor binding sites, suggesting a mechanism other than a ternary complex model for enhanced occupancy (Heinz et al. 2010). Gain-of-function studies indicated rapid nucleosome remodeling following collaborative binding of PU.1 to the regions destined to acquire enhancer-like features in myeloid cells based on subsequent H3K4 mono-methylation (Heinz et al. 2010). Furthermore, PU.1 is required for the maintenance of the macrophage epigenome and its expression in PU.1-negative myeloid progenitors is sufficient to induce nucleosome remodeling and H3K4 mono-methylation at cell-specific sites (Heinz et al. 2010).

SDTFs bind in a cell-specific manner

Genome-wide studies to date of SDTFs indicate that while binding at promoters is enriched over that expected by chance, the majority of binding sites for these factors are located at inter- and intra-genic locations that exhibit the features of enhancers. Initially demonstrated for the estrogen receptor (Lupien et al. 2008), this pattern of genome-wide binding is also observed for other members of the nuclear receptor family, including the glucocorticoid receptor (Biddie et al. 2011), LXR (Heinz et al. 2010), PPARG (Lefterova et al. 2008), androgen receptor (AR; Lupien et al. 2008, He et al. 2010, Wang et al. 2011), estrogen receptor-related receptors (Lupien et al. 2008, Hurtado et al. 2011, Li et al. 2013), and nuclear receptor subfamily 1 group D members 1 and 2 (NR1D1 and NR1D2), known colloquially as Rev-Erb (Lam et al. 2013). Similarly, studies on transcription factors responsive to extracellular signals also primarily localize to distal enhancer-like elements, including the RELA component of NF-κB (Barish et al. 2010, Kaikkonen et al. 2013), serum response factor (SRF; Sullivan et al. 2011), FOXO1 (Fan et al. 2010), STAT1 (Vahedi et al. 2012), and SMAD3 (Mullen et al. 2011). Although identification of specific enhancer–promoter interactions remain challenging at a genome-wide level, the binding of SDTFs at distal locations is correlated with changes in expression of nearby genes (Kaikkonen et al. 2013, Lam et al. 2013).

In all of the cases in which the genome-wide locations of SDTFs have been examined in macrophages, a large fraction of the DNA-binding sites are observed to be in close proximity to the binding sites for one or more macrophage LDTFs (PU.1, C/EBPs, and/or AP1 factors) (Heinz et al. 2010). Because LDTFs select enhancers in a cell-specific manner, the colocalization of SDTFs implies a cell-specific binding pattern. Consistent with this, the genomic locations of PPARG and Rev-Erbbs in macrophages are very different from their genomic locations in adipocytes (Lefterova et al. 2008) and hepatocytes (Cho et al. 2012) respectively.

An important question is the hierarchical relationship between LDTFs and SDTFs, as well as their respective roles in the selection of cell-specific enhancers. Studies of LDTFs and SDTFs in macrophages suggest the primary roles of LDTFs, with variable roles of SDTFs that depend on the specific factor analyzed. Genome-wide binding analyses of LXR indicated that nearly all LXR-binding sites were in close proximity to the binding sites for macrophage LDTFs (PU.1, API, C/EBPs) (Heinz et al. 2010). Genetic deletion of LXR had virtually no effect on the overall enhancer atlas in macrophages and did not result in a loss of nearby PU.1 binding (Heinz et al. 2010). In contrast, deletion of PU.1 resulted in loss of nearby LXR binding (Heinz et al. 2010). Thus, in this case, a clear hierarchy was observed in which the binding of LXRs was dependent on the initial selection of enhancers by LDTFs.

A different picture emerged from studies of the RELA subunit of NF-κB, which is strongly activated in macrophages by signaling through TLR4. Following activation and nuclear entry, ~85% of RELA binding occurred at pre-existing enhancer-like regions of the genome that were occupied by macrophage LDTFs (Barish et al. 2010, Escoubet-Lozach et al. 2011, Kaikkonen et al. 2013). However, ~15% of the binding sites for RELA occurred at the genomic regions that lacked features of enhancers, but acquired these features subsequent to RELA binding (Escoubet-Lozach et al. 2011, Kaikkonen et al. 2013). Intriguingly, these binding sites were highly enriched for nearby recognition motifs for the macrophage LDTFs API and PU.1. Subsequent experiments demonstrated that RELA binding required PU.1, while PU.1 binding required nuclear entry of RELA (Escoubet-Lozach et al. 2011, Kaikkonen et al. 2013). Thus, in this case, RELA acted as both a signal-dependent and collaborating transcription factor to open up regions of chromatin and select new enhancers. The ability of RELA but not LXR to perform this function may relate to relatively higher nuclear concentrations and/or its ability to interact with nucleosome-remodeling factors.

Recent studies have utilized natural genetic variation as an in vivo mutagenesis screen to examine the collaborative binding model for LDTFs and the hierarchical relationship with NF-κB. Genome-wide transcription factor binding and gene expression analyses were used to characterize how genetic variations affect the binding of
LTDFs (PU.1 and CEBPA) and the associated transcriptional output in macrophages derived from two inbred rodent strains, C57BL/6j and BALB/c (Heinz et al. 2013). Single nucleotide variations and other noncoding genomic variants were shown to directly perturb LTDF as well as SDTF binding (Heinz et al. 2013). However, these studies have also demonstrated that a variant in a binding motif for an LTDF not only affected the binding of that factor but also the binding of nearby LTDFs, even though there were no alterations in the binding motifs for those factors. In the case of differential NF-kB binding in the two mouse strains, mutations in motifs of LTDFs were three times more likely to result in decreased RELA binding than mutations in the binding sites for RELA itself (Heinz et al. 2013). These findings provided genetic evidence that collaborative binding of LTDFs to variably spaced sequences was essential for enhancer selection and subsequent acquisition of histone modifications associated with enhancer activity (Heinz et al. 2013). These studies also provided genetic evidence for a hierarchical relationship between LTDFs and the majority of the binding sites for NF-kB.

From these studies, the authors also formulated a genetic definition of a functional binding matrix that could be used to predict whether a specific variant in an LTDF-binding motif would elicit a functional consequence. These results hint at the potential power of defining functionally important noncoding regulatory variants and suggest a strategy for prioritizing natural genetic variants that may play roles in human disease (Heinz et al. 2013).

A similar study detailed the binding profile of three tissue specific transcription factors (HNF4A, CEBPA, and FOXA1) in the livers of six inbred rodents (Stefflova et al. 2013). From these studies, the authors describe that although there are extensive transcription factor binding differences among relatively similar species, single nucleotide variation plays only a modest role in these differences (Stefflova et al. 2013). Regions that are bound by combinations of LTDFs are more evolutionarily stable; however, binding by a single LTDF can become destabilized through the deletion of one of its LTDF co-binding partners (Stefflova et al. 2013).

Lineage-determining factors in other cell types

The roles of LTDFs in mediating cell-specific and signal dependent activation of transcription have been described in many cell types. Here, we provide examples derived from genome-wide studies in dendritic cells, T cells, and in model systems used to study transforming growth factor β (TGFβ)-mediated and hormone-dependent gene expression (Table 2).

Dendritic cells

Dendritic cells and macrophages derive from a common myeloid lineage. In support of a LTDF/SDTF regulatory model, H3K4me3, H3K4me1, H3K27Ac, and a comprehensive panel of myeloid transcription factors were utilized to map the promoter and enhancer landscape in dendritic cells (Garber et al. 2012). Like in macrophages, the LTDFs in dendritic cells include PU.1 and CEBPB (Garber et al. 2012). In this model, a second tier of LTDFs (JUNB, IRF4, and ATF3) exists, termed ‘primers’, which co-bind with PU.1 and CEBPB at LPS-inducible genes pre-stimulation. Later, these factors become associated with specific and dynamic activation factors, mediating a signal-specific response to stimulation (Garber et al. 2012).

T cells

Three seminal papers have described the cell-specific enhancer landscape in T cells, focusing on variations amongst different subsets of T helper (Th) cells. In one study, the authors sought to define the subset-specific enhancer landscape in regulatory T (Treg) cells. LTDFs in the RUNX (Runt-related transcription factor) and ETS families, CBFB, ETS1, and ELF1, were shown to predominantly occupy quiescent Th cell enhancers (Samstein et al. 2012). Surprisingly, the majority of FOXP3 binds to both already established and Treg-polarized enhancers (Samstein et al. 2012). Only 2% of FOXP3-binding sites appear to be Treg-specific and are highly enriched for an AP1 motif (Samstein et al. 2012). In some cases, FOXP3 was found to displace its paralog, FOXO1, or other cofactors that served as placeholders for FOXP3 binding (Samstein et al. 2012). Furthermore, during Treg polarization and TCR activation, the SDTF, FOXP3, is primarily recruited to previously established enhancers likely through AP1 and nuclear factor of activated T-cells (NFAT) facilitated chromatin remodeling (Samstein et al. 2012). A second study describing the transcriptional network for Th17 polarization has found two previously established LTDFs, BATF and IRF4, to exhibit similar binding patterns between quiescent Th cells and Th17 cells (Ciofani et al. 2012). In the presence of Th17-polarizing cytokines, BATF and IRF4 prime pre-established enhancers for SDTF (RORG, STAT3, and MAF) recruitment (Ciofani et al. 2012).
Lastly, to define the active enhancer landscape after Th1 or Th2 polarization, a comprehensive analysis of EP300 binding was conducted in the two subsets (Vahedi et al. 2012). The two closely related cell subsets display common and distinct EP300 binding (Vahedi et al. 2012). Moreover, STAT6 and GATA3 or STAT4/STAT1 and TBX21 were found to activate Th2 or Th1 subset-specific enhancers, respectively, while suppressing those of alternative fates (Vahedi et al. 2012).

Embryonic stem cells, pro-B cells, myotubes

Genome-wide binding analyses were conducted for three unique cell types: embryonic stem cells, myotubes, and pro-B cells, each in the quiescent state and after TGF-β signaling (Mullen et al. 2011). The authors found that the LDTFs for these respective cell types, POU5F1, MYOD1, and PU.1, directed the TGFβ-induced SDTF, SMAD3, to unique cell-specific enhancer sites (Mullen et al. 2011). Motif-finding at SMAD3 bound sites showed an enrichment for LDTF motifs specific to the cell type in question, highlighting the importance of LDTFs in directing the binding of SDTFs (Mullen et al. 2011).

Breast, prostate, and osteoblasts

FOXA1 is essential for the development and differentiation of several organs such as liver, kidney, pancreas, lung, prostate, and mammary gland (Friedman & Kaestner 2006). In breast cancer epithelial (MCF7) cells and prostate cancer (LNCaP) cells, FOXA1 creates a cell-specific enhancer network, which regulates other activating factors such as ESR1 (Tan et al. 2008, Hurtado et al. 2011). In estrogen-responsive breast cancer cells, tamoxifen-mediated ESR1 activity is dependent on LDTF FOXA1. In tamoxifen-resistant cells, ESR1 binding occurs independent of ligand but is still dependent on FOXA1 (Hurtado et al. 2011). Another transcription factor implicated in breast cancer oncogenesis, TFAP2C, binds to ESR1-binding sites in a ligand-independent manner, colocalizing with FOXA1 and priming enhancer sequences for ESR1 binding (Tan et al. 2011). Furthermore, treatment with estradiol results in a global increase in enhancer RNA (eRNA) transcription at enhancers in the vicinity of 17β-estradiol-regulated genes, which promote specific E2/ESR1/eRNA-induced enhancer-promoter looping

Table 2  Summary of studies demonstrating signal-dependent cell-specificity in transcription

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Lineage-determining transcription factors</th>
<th>Stimulus</th>
<th>Stimulus-activated transcription factors</th>
<th>Cell-specific (LDTF-regulated) gene expression</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macrophages</td>
<td>PU.1, C/EBPs, AP1</td>
<td>LPS, GW3965 (LXR agonist)</td>
<td>LXRα, RELA, BCL6</td>
<td>Csf1r, Cd14, Ccl2, Ccl4, Cxcl12, Abcg1</td>
<td>Barish et al. (2010) and Heinz et al. (2010)</td>
</tr>
<tr>
<td>Dendritic cells</td>
<td>PU.1, CEBPB</td>
<td>LPS</td>
<td>JUNB, IRF4, ATF3</td>
<td>Il1a, Ifit1, Ifit3, Tnf, Stat1, Nfkbia</td>
<td>Garber et al. (2012)</td>
</tr>
<tr>
<td>ES cells</td>
<td>POU5F1, PU.1, MYOD1</td>
<td>TGFB1</td>
<td>SMAD3</td>
<td>Sox2, Vpreb2</td>
<td>Mullen et al. (2011)</td>
</tr>
<tr>
<td>Pro-B cells</td>
<td>BATF and IRF4</td>
<td>Th17 polarization</td>
<td>SMAD3</td>
<td>Adora1</td>
<td>Mullen et al. (2011)</td>
</tr>
<tr>
<td>T cells</td>
<td>?</td>
<td>Th1 polarization</td>
<td>SMAD3, ROR, STAT4, STAT1, TBX21</td>
<td>Il17a, Il12rb1</td>
<td>Mullen et al. (2011)</td>
</tr>
<tr>
<td>T cells</td>
<td>CBFB, ETS1, ELF1</td>
<td>Th2 polarization</td>
<td>STAT6, GATA3</td>
<td>Junb</td>
<td>Vahedi et al. (2012)</td>
</tr>
<tr>
<td>MCF7 cells</td>
<td>FOXA1, TFAP2C</td>
<td>Tamoxifen, 17β-estradiol</td>
<td>ESR1</td>
<td>Ret, Xbp1, Tff1</td>
<td>Samstein et al. (2012)</td>
</tr>
<tr>
<td>LNCaP cells</td>
<td>FOXA1</td>
<td>5α-dihydrotestosterone</td>
<td>AR</td>
<td>Chka</td>
<td>Lupien et al. (2008), Hurtado et al. (2011), Tan et al. (2011) and Li et al. (2013)</td>
</tr>
<tr>
<td>Murine mammary epithelial cells</td>
<td>AP1</td>
<td>Dexamethasone</td>
<td>Glucocorticoid receptor</td>
<td>Klb1c</td>
<td>Lupien et al. (2008), He et al. (2010) and Wang et al. (2011)</td>
</tr>
<tr>
<td>U2OS</td>
<td>GATA4</td>
<td>17β-estradiol</td>
<td>ESR1</td>
<td>Fasl, Wnt4, Foxc1, Alpl, Runx2</td>
<td>Biddie et al. (2011)</td>
</tr>
</tbody>
</table>

Cell-specific functions of signal-dependent transcription factors

Thematic Review

D X Zhang and C K Glass

DOI: 10.1530/JME-13-0216

Published by Bioscientifica Ltd

Downloaded from Bioscientifica.com at 11/17/2018 12:26:15AM via free access
In addition, a recent study conducted in a murine mammary epithelial cell line has found AP1 to be an essential LDTF, mediating GR-chromatin interactions and GR-regulated transcription (Biddie et al. 2011).

In U2OS cells, an osteoblast-like cell line, ESR1 binding is preceded by the lineage determining factor, GATA4 (Miranda-Carboni et al. 2011). GATA4 is necessary for H3K4me2 deposition at ESR1-binding sites, suggesting that it is a LDTF for ESR1 recruitment in this cell type (Miranda-Carboni et al. 2011). Furthermore, these results illustrate the cell-specific nature of priming and activating transcription factors; future studies will help clarify the differences amongst diverse cell types.

**Histone modifications associated with active and poised regulatory elements**

LDTFs and SDTFs bind at both enhancer and promoter elements to regulate transcription. Enhancers and promoters are associated with distinct chromatin signatures – active promoters are marked by high levels of histone H3 lysine 4 tri-methylation (H3K4me3) relative to mono-methylation (H3K4me1), whereas enhancers are marked by high levels of H3K4me1 relative to H3K4me3 (Heintzman et al. 2007; Fig. 2). H3K4me2 occurs at both promoter and enhancer regions (Kaikkonen et al. 2013), which we define here by their distance from a transcription start site (around \(<2\,\text{kb}\) to indicate a promoter and \(>2\,\text{kb}\) to indicate an enhancer). Studies in yeast, *Drosophila*, and humans suggest that the SET domain-containing methyl-transferases are responsible for depositing the majority of H3K4me3 at promoters (Nagy et al. 2002, Wu et al. 2008, Ardehali et al. 2011, Hallson et al. 2012).

Mono-methylation at enhancers is dynamic and regulated in a cell-specific manner. Currently, methyl-transferases regulating H3K4me1 at mammalian enhancers include MLL1 (Jeong et al. 2011) and MLL1/3 (Kaikkonen et al. 2013). At enhancer loci, H3K4me1 functions as an active mark; its de-methylation in mouse embryonic stem cells by LSD1 results in enhancer silencing and cell differentiation. The process by which H3K4me1 is lost at enhancers is called enhancer 'decommissioning' and further illustrates the fine tuning mediated by cell-specific enhancers (Whyte et al. 2012). Not surprisingly, loss of LSD1 is associated with increased H3K4me1 and H3K4me2 at hematopoietic stem cell promoter and enhancer elements, as well as subsequent gene derepression, resulting in altered transcription and compromised differentiation (Kerenyi et al. 2013).

While the presence of H3K4me2 dictates active transcriptional regions, active enhancers are also marked

![Selected histone marks at promoters and enhancers](http://jme.endocrinology-journals.org)

**Figure 2**

Histone marks at active and poised promoters and enhancers. Promoters are designated by high levels of H3K4me3, while enhancers contain high levels of H3K4me1/2. In the poised or repressed state, promoters contain histone methylation marks, H3K27me3, and H3K20me3, while enhancers contain H3K27me3/me1 and H3K9me1. Active promoters and enhancers are marked by H3K4me2 and H3K27Ac. GTF, general transcription factor; LDTF, lineage-determining transcription factor; SDTF, signal-dependent transcription factor.

In contrast to active acetylation marks, the combination of H3K4me3 and H3K27me3 marks promoters that are ‘poised’ for transcription (Mikkelsen et al. 2007; Fig. 2). In mouse and human embryonic stem cells, poised enhancers are marked by H3K4me1, EP300, SMARCA4, and H3K27me3, and are bound by the Polycomb complex (Rada-Iglesias et al. 2011). Unable to drive gene expression, poised enhancers transition into active enhancers during differentiation through the acquisition of H3K27ac and the loss of H3K27me3, as well as the recruitment of tissue-specific transcription factors and RNA polymerase II (Rada-Iglesias et al. 2011). Furthermore, poised enhancers in hematopoietic stem cells contain H3K27me1 and H3K9me1 before their activation (Cui et al. 2009), whereas promoters in embryonic stem cells (Rada-Iglesias et al. 2011), hematopoietic stem cells (Cui et al. 2009), and adult tissues (De Santa et al. 2007), all contain H3K27me3 and may contain other repressive marks such as H3K20me3 (Stender et al. 2012; Fig. 2).

Signal-dependent control of the epigenetic landscape

Investigation of macrophage differentiation and activation has provided a powerful experimental system for linking the actions of LITFs and SDTFs to specific histone modifications and subsequent transcriptional output. Macrophage gene expression is stimulus-specific, genespecific, cell-specific, and macrophage-subset specific. In early genome-wide studies to study the stimulus-specific nature of gene expression, microarrays were used to systematically examine the response of macrophages to various bacterial pathogens that act through TLRs, comparing the shared and differential transcriptional output (Nau et al. 2002). These studies were the first to document the very dramatic transcriptional responses of macrophages to these signals and revealed that these responses were finely tuned to specific pathogens.

Several lines of evidence indicate that the program of macrophage activation is subject to many types of negative regulation that act prior and/or subsequent to stimulus responses. BCL6 has been shown to corepress almost a third of the TLR4 cistrome, such that its loss results in hypersensitivity to pro-inflammatory derepression (Barish et al. 2010). The NCOR1/HDAC3 corepressor complex has been proposed to function as a transcriptional checkpoint for some TLR-responsive genes under basal conditions through recruitment to non-phosphorylated forms of AP1. The HDAC3 component of this complex contributes to the repressive functions of the NCOR1 complex by removing histone acetylation marks required for transcriptional activity (Ogawa et al. 2004; Fig. 3). Activation of inflammatory genes necessitates signal-dependent phosphorylation of c-Jun and removal of NCOR1 from AP1 target genes (Ogawa et al. 2004). Surprisingly, loss of HDAC3 in macrophages results in decreased activation of almost half of the inflammatory program (Chen et al. 2012). This phenomenon has been suggested to be the result of decreased Ifnb1 expression and the secondary STAT1-mediated transcriptional response, a pre-requisite for inflammatory induction (Chen et al. 2012). In addition to HDAC3, NCOR1 complexes in macrophages also contain the histone methyl-transferase SMYD5, which contributes to repression by catalyzing H4K20 tri-methylation, inhibiting the expression of TLR4 target genes (Stender et al. 2012). Signal-dependent de-methylation of H4K20me3 at promoters occurs through the recruitment of the histone de-methylase PHF2 by the RELA component of NF-κB (Stender et al. 2012; Fig. 3).

After corepressor dismissal from inflammatory genes, transcriptional activation in response to TLR4 signaling occurs in two phases – primary response genes are immediately upregulated in response to stimuli whereas secondary response genes require expression of gene products from the initial wave of transcription, such as the type I interferons. CpG islands are usually found at the promoters of primary response genes and associate with nucleosome-destabilizing elements, allowing these genes to be transcribed at low levels in the basal state. This low level of transcription produces nonfunctional transcripts that are rapidly degraded until the recruitment of the positive transcription elongation factor (P-TEFb) complex, which phosphorylates RNA polymerase II and couples transcriptional initiation with elongation (Hargreaves et al. 2009, Ramirez-Carrozzi et al. 2009). In contrast, promoters with low CpG content are indicative of secondary response genes, which undergo stimulus-dependent
H3K4me3 and H3 acetylation, requiring selective nucleosome remodeling (Hargreaves et al. 2009, Ramirez-Carrozzi et al. 2009). Inflammatory genes in macrophages can be further classified into two categories on the basis of function and regulation – genes capable of being induced after stimulation of tolerant macrophages and those that will not be induced during restimulation. Genes capable of being restimulated exhibit more H4 acetylation and maintain H3K4me3 after restimulation (Foster et al. 2007). Like the genes in naïve macrophages, these genes are capable of recruiting the SMARCA4 and CHD4 chromatin-remodeling complexes to their promoters; however, they exhibit different transcriptional requirements than those in naïve macrophages (Foster et al. 2007).

Macrophage activation results in the sequential binding of lineage-determining and stimulus-activated transcription factors to cell-specific regions, enabling the novel deposition of cis-regulatory enhancer marks such as H3K4me1 and H3K27ac (Escoubet-Lozach et al. 2011, Kaikkonen et al. 2013, Ostuni et al. 2013). In one study, the authors removed the stimuli post-treatment in a ‘washout’ experiment. Surprisingly, H3K4me1 did not revert to its latent state; instead, it persisted and conditioned a faster and stronger response upon restimulation, suggesting the existence of epigenetic memory in response to stimuli exposure (Ostuni et al. 2013).

Recent studies have reported the productive transcription of RNA polymerase II-associated noncoding RNA (De Santa et al. 2010) (ncRNA) from cis-regulatory enhancers located both intragenically (Kowalczyk et al. 2012) and intergenically in response to LPS stimulation (Kaikkonen et al. 2013). TLR4 signaling regulates macrophage gene expression through both a pre-existing enhancer landscape as well as the induction of ∼3000 eRNAs from de novo enhancer regulatory regions (de novo eRNAs) (Kaikkonen et al. 2013; Fig. 1). In all, ∼2200 de novo eRNAs are induced greater than twofold in response to TLR4. Regions that show eRNA induction also gain H3K4me2 and are enriched for AP1, NF-κB, C/EBPs,
IRF/STAT, and PU.1 motifs (Kaikkonen et al. 2013). Surprisingly, the inhibition of RNA polymerase II elongation and eRNA synthesis at enhancers resulted in decreased H3K4me1 and H3K4me2 deposition without affecting H4K8ac (Kaikkonen et al. 2013). This result suggests that enhancer transcription precedes and may be essential for H3K4me1/2 deposition, at least for the set of enhancers that are established de novo following TLR4 activation. Systemic depletion of all known H3K4 methyl-transferases indicated that TLR4-induced H3K4 methylation was dependent upon the histone methyl-transferases, MLL1, MLL2/4, and MLL3 (Kaikkonen et al. 2013).

**Functional roles of eRNAs**

It has been proposed that transcription at enhancers maybe due to nonspecific interactions of RNA polymerase II with the genome, thus representing noise rather than biological function (Struhl 2007). Others have speculated that the functional effects of enhancer transcription are due to the transcription process and machinery or RNA polymerase II moving unobstructed in either direction upon binding DNA. In contrast, studies have shown that in vivo developmental enhancers have highly conserved noncoding elements (Pennacchio et al. 2006, Visel et al. 2008), positing that developmental enhancers may have retained their function throughout evolution.

eRNAs are believed to exert their function through two main mechanisms: in cis, the eRNAs may act on the same chromatin fiber or in trans, the eRNAs may function at distant sites in the genome. Recently, in macrophages, Rev-Erbα (NR1D1 and NR1D2, more commonly referred to as Rev-Erbα, see Table 1) have been described to repress key inflammatory genes through direct promoter-mediated mechanisms as well as in cis through eRNAs (Lam et al. 2013). Primarily, Rev-Erbα bind at distal elements, which are marked by PU.1 co-binding and H3K4me1 (Lam et al. 2013). The binding of Rev-Erbα at enhancer elements inhibits enhancer mediated transcription at two distal cis regulatory elements, subsequently decreasing transcription at the nearby protein-coding genes, Mmp9 and Cx3cr1 (Lam et al. 2013). Genetic studies depleting either Rev-Erbα from these loci or the eRNAs themselves resulted in the subsequent derepression of Mmp9 and Cx3cr1 mRNA (Lam et al. 2013). Similar studies on multiple cell types have found the transcription of ncRNAs, especially eRNAs, to be essential in promoting (Feng et al. 2006, Kim et al. 2010, Orom et al. 2010, Li et al. 2013, Melo et al. 2013) and inhibiting (Wang et al. 2008) gene expression. Further studies will help to elucidate the cell-specific mechanisms of enhancers in regulating diverse facets of transcription.

**Conclusions**

In concert, these studies suggest a general model to explain how genes with a broad distribution of expression can be regulated in a cell-specific manner and how broadly expressed SDTFs are capable of regulating discriminative responses in different cell types. Traditionally, the regulation of transcription is thought to occur primarily at the promoter. The above-described studies show that LDTFs collaboratively select and prime distal enhancer regulatory elements in a cell-specific manner, genome-wide. In a signal-dependent manner or during development, LDTFs may remodel and open chromatin at promoter and enhancer regions, broadly defining the regulatory potential of genes. For example, PU.1 binds to more than two-thirds of enhancer-like regions in macrophages, priming target genes for the subsequent recruitment of SDTFs to their vicinity, and their subsequent activation (Heinz et al. 2010). Thus, although transcription can occur at both promoters and enhancers – transcription at promoters is often ubiquitous amongst different cell-types while that at enhancers tends to be cell-specific.

Furthermore, genetic analyses have shown that mutations of SDTF motifs can only account for a small percentage of the variable gene expression that is the result of genetic variation (Kasowski et al. 2010, Heinz et al. 2013). Understanding how mutations in LDTF and SDTF motifs result in direct and indirect effects on enhancer selection and function is essential to defining relevant noncoding variants in the genome and their impact on human disease (Heinz et al. 2013). Comprehension of how key lineage determining transcription factors modulate signal-dependent mechanisms in different cell types may also have cell-specific therapeutic applications. Many human disease states could benefit from cell-specific alterations in gene expression in a manner that would decrease toxicity and increase efficacy. New technologies such as antisense oligonucleotides and small interfering RNAs can be potentially used to modulate cell-specific eRNA transcripts that are associated with malfunctioning genes (Lam et al. 2013). In principle, this methodology would enable the cell-specific regulation of aberrant gene expression implicated in disease states such as cancer and chronic inflammatory diseases without adversely modulating normal physiological expression in unaffected cell types.


Received in final form 11 October 2013
Accepted 15 October 2013
Accepted Preprint published online 15 October 2013