Deciphering *cis*-regulatory control in inflammatory cells

Serena Ghisletti and Gioacchino Natoli

Department of Experimental Oncology, European Institute of Oncology (IEO), Via Adamello 16, 20139 Milan, Italy

In innate immune system cells, such as macrophages and dendritic cells, deployment of inducible gene expression programmes in response to microbes and danger signals requires highly precise regulatory mechanisms. The inflammatory response has to be tailored based on both the triggering stimulus and its dose, and it has to be unfolded in a kinetically complex manner that suits the different phases of the inflammatory process. Genomic characterization of regulatory elements in this context indicated that transcriptional regulators involved in macrophage specification act as pioneer transcription factors (TFs) that generate regions of open chromatin that enable the recruitment of TFs activated in response to external inputs. Therefore, competence for responses to a specific stimulus is programmed at an early stage of differentiation by factors involved in lineage commitment and maintenance of cell identity, which are responsible for the organization of a cell-type-specific *cis*-regulatory repertoire. The basic functional and organizational principles that regulate inflammatory gene expression in professional cells of the innate immune system provide general paradigms on the interplay between differentiation and environmental responses.

1. Introduction

Transcription of eukaryotic genes is a highly coordinated process that is regulated in time and space. At the heart of transcriptional control is the vast amount of DNA-encoded regulatory information contained in mammalian genomes and located both right upstream of transcription start sites and at distal regulatory elements, particularly enhancers.

Cell-type-specific usage of the vast genomic repertoire of *cis*-regulatory sequences is accomplished mainly by the combinatorial activity of sequence-specific transcription factors (TFs) involved in lineage determination and maintenance of cell identity [1]. Genomic regions that are active as enhancers in a given cell type typically reside in regions with altered chromatin structure, as revealed by hypersensitivity to DNase enzymes and nucleosomal depletion [2–5]. In addition, genome-wide studies indicate that enhancers exhibit a characteristic chromatin ‘signature’, consisting of monomethylation of lysine 4 in histone H3 (H3K4me1) in the absence of significant trimethylation (H3K4me3) [6–9]. Additional marks associated with active enhancers include binding of histone acetyltransferases such as p300 and CBP, and histone acetylation (most notably, but not at all exclusively H3K27) [10–13]. High-resolution maps of these histone modifications indicated that the regions of the genome bearing a chromatin signature of enhancers are generated in a cell-type-specific manner [12,14–17]. According to current models, the cell-type specificity of these regulatory regions is due to the presence of binding sites for cell-type-specific TFs that are directly implicated in their functionalization. Evidence accumulated during the past few years suggests that lineage-specific TFs organize the enhancer landscape required for differential recruitment of TFs involved in gene expression programmes during development or in response to external stimuli [18–22]. Although this model implies that the enhancer repertoire in a given cell type is predetermined, the emerging picture indicates that chromatin modifications at these genomic regulatory elements can be highly dynamic in order to allow cell-type-specific epigenomes to acquire their characteristic properties or to respond
to signals received from the microenvironment. In this context, the epigenome of cells whose main biological function is to sense and to react to environmental inputs, such as cells of the innate immune system, has to rapidly change according to the received signals. This review will focus on the epigenomic landscape of inflammatory cells as a paradigmatic example of how the genome can be organized to enable highly specialized responses to a broad array of external stimuli.

2. The epigenome of inflammatory cells

The innate immune system provides initial protection against infection, injury and tissue stress. Genes encoding proteins with antimicrobial and proinflammatory activities must be rapidly and highly induced in the presence of an invading microbe as well as upon release of the intracellular content of damaged tissues. At the same time, inflammatory genes must be maintained in a transcriptionally repressed state under normal homeostatic conditions.

(a) General features of the inflammatory gene expression programme

The complexity of the inflammatory response requires several hundreds genes to be activated in a kinetically complex fashion, with some genes rapidly activated immediately after the stimulus (‘primary’ response genes) and others induced with slower kinetics (‘secondary’ response genes and some slowly activated primary response genes). Regardless of their activation kinetics, primary response genes are formally defined as those genes that can be induced without de novo protein synthesis, while secondary genes require new protein synthesis [23]. Moreover, some genes, such as the one encoding TNFα, are activated with similar patterns in almost every type of inflammatory responses while others (that encode proteins with more specialized functions) are tightly regulated in a stimulus-specific and cell-type-restricted manner [24]. During an inflammatory response, a large number of genes are coordinately induced by a limited set of TFs that are constitutively expressed by many cell types, such as NF-κB, IFN-regulatory factors (IRFs) and AP-1 family members. In spite of the involvement of these general, broadly expressed regulators, inflammatory gene expression is both cell-type- and stimulus-specific, an observation that until recently did not have an adequate mechanistic explanation. Clearly, the transcriptional selectivity has to be established by other regulatory layers.

First of all, analysis of the genomic features of the promoters of primary response genes revealed a striking enrichment for CpG islands, a sequence feature that directly or indirectly bias their expression programmes in primary response genes. This suggests the existence of a ‘basal state’ for the genome of inflammatory cells, a state characterized by an organized epigenetic landscape that ensures the expression of the vast majority of primary response genes in a coordinated fashion, at the expense of secondary response genes.

(b) Organization and function of cis-regulatory regions in inflammatory cells

Recent studies on macrophages, cells that play a central role in the inflammatory response, link changes in chromatin structure to molecular events occurring during lineage commitment and development that provide competence for subsequent transcriptional induction. ChIP-Seq studies revealed that macrophages undergo a large number of histone modifications that are characteristic of active chromatin [25]. In addition, primary response genes associated with CpG islands are constitutively associated with RNA polymerase II and transcribed at low levels [26]. Overall, it is clear that a different sequence-programmed chromatin organization at individual inflammatory genes lays the grounds for different expression kinetics.

3. Pioneer transcription factors and the control of cis-regulatory regions

Recent technological advances have started to elucidate the timing and mechanisms by which specific enhancers acquire unique chromatin signatures. Evidence accumulated during the last few years indicates that during development some
TFs start marking tissue-specific regulatory elements. Pioneer factors are functionally defined as sequence-specific DNA-binding proteins able to bind to their target sites when embedded in a nucleosomal context that is not permisive for binding of other TFs. Subsequent recruitment of chromatin remodelers by pioneer factors results in stable local opening of the chromatin, thus making it competent for other factors to bind (figure 1). Once bound, pioneer factors act in some cases as placeholders that will be replaced by other TFs at later stages of development [33]. In this context, a relay model, involving the replacement of pluripotent factors at some primed regions by related lineage-specific TFs has been proposed, such as the replacement of Sox2 by Sox4 on primed regions during B-cell differentiation [34]. Also in completely differentiated cells, the emerging picture is that pioneer TFs can access their binding sites even when wrapped in nucleosomes. Consistent with this model, these lineage-specific TFs are actively involved in determining the baseline accessible chromatin landscape in order to facilitate the recruitment of other TFs unable to invade nucleosomes, including many of those responsive to environmental stimuli.

(a) Role of PU.1 as a global genome organizer in inflammatory cells

In macrophages, the observations reported above suggest a model wherein the global landscape determined by Pu.1 provides the context in which the dynamic TF–chromatin interactions in response to specific environmental inputs occur. Pu.1 expression in non-myeloid cells or in Pu.1-negative myeloid progenitors is sufficient to induce nucleosome-free DNA stretches limited on both sides by H3K4me1-marked nucleosomes at the same genomic regions identified as enhancers in macrophages [10]. Therefore, these data clearly indicate that Pu.1 is not a simple marker of enhancers but indeed controls formation and accessibility of the entire genomic regulatory repertoire of macrophages. Such activity implies that Pu.1 must be able to bind nucleosomal sites, attracting chromatin-remodelling factors to stabilize nucleosome

(b) Dynamic changes of cis-regulatory regions in response to stimuli

So far, several studies have reported that dynamic changes in chromatin marks correlate with transcriptional activity. H3K4me1 represents a general mark of distal regulatory regions and additional modifications can distinguish between enhancers that are active and those that are poised and can subsequently be activated during developmental transitions or in response to external stimuli. Recent findings pointed out that only a fraction of distal H3K4me1-positive regions involved in modulating transcription in a given cell type are also marked by H3K27 acetylation [43–45]. These enhancers are defined as active, as opposed to inactive and poised enhancers containing H3K4me1 only [43,44]. Further computational analyses defined additional sub-categories of active (H3K4me1/H3K27Ac positive), poised (H3K4me1/H3K27me3 positive), or intermediate (H3K4me1 positive/H3K27Ac negative) enhancers using H3K27Ac in combination with other chromatin features [46].
signalling pathway. In a paradigmatic case, the cell-type-specific master regulator is responsible for determining the cell-type-specific effects of the same stimulus is able to activate a distinct enhancers repertoire in different cells has been described. From a conceptual point of view, the general principle is that master regulators are responsible for driving developmental transitions are associated with dynamic changes in the H3K4me1 repertoire. Upon commitment to the B or myeloid lineages, an increase in H3K4me1 was observed across genomic regions associated with a B-cell or myeloid lineage gene expression programme, respectively. Thus, developmental progression from the multipotent haematopoietic progenitor to a committed lymphoid or myeloid cell stage, lineage-restricted enhancer elements were found to be already primed by H3K4me1 prior to differentiation [47]. More importantly, the mechanism by which the same stimulus is able to activate a distinct enhancers repertoire in different cells has been described. From a conceptual point of view, the general principle is that master regulators are responsible for determining the cell-type-specific effects of the same signalling pathway. In a paradigmatic case, the cell-type-specific effects of TGFβ signalling were found to be determined by the interaction of the TGFβ-activated TFs Smad2/3 with TFs that specify and maintain cell identity [42]. The genomic distribution of Smad3 was found to overlap Oct4 in embryonic stem cells (ESCs), Myod1 in myotubes and Pu.1 in pro-B cells. These cell-type-specific master regulators allow Smad3 binding by establishing open chromatin accessible regions first and by tethering Smad3 to them via protein–protein interactions when the signal is delivered [42].

In the innate immune system, external inputs such as proinflammatory stimuli are expected to have a broad impact on chromatin organization and marking at cis-regulatory elements. In macrophages, a recent genome-wide analysis confirmed that thousands of cis-regulatory regions gain H4 acetylation upon the proinflammatory stimulus lipopolysaccharide (LPS) as expected [48]. These studies also suggested a specific and non-redundant role of the histone deacetylase HDAC3 in controlling acetylation levels at a subset of genomic regions. In this context, it would be of interest to study how distinct stimuli can reorganize the enhancer repertoire of professional innate immune cells to clearly define general principles regulating the dynamic use of the available cis-regulatory information. It is possible that the impact of a broad array of stimuli on such highly specialized cells goes far beyond simple changes in acetylation and results in a reorganization of a subset of enhancers.

### Table 1. Summary of studies on master regulators that generate regions of open chromatin enabling the recruitment of stimulus-activated TFs in different cell types.

<table>
<thead>
<tr>
<th>cell type</th>
<th>master regulator</th>
<th>stimulus</th>
<th>stimulus-activated TFs</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>dendritic cells</td>
<td>Pu.1</td>
<td>LPS</td>
<td>NF-κB, STATs, AP-1, IRFs</td>
<td>[36]</td>
</tr>
<tr>
<td>LNCaP (prostate epithelium)</td>
<td>FoxA1</td>
<td>5α-dihydrotestosterone</td>
<td>AR</td>
<td>[37]</td>
</tr>
<tr>
<td>MCF-7 (breast epithelium)</td>
<td>FoxA1</td>
<td>tamoxifen, oestradiol</td>
<td>ERα</td>
<td>[38]</td>
</tr>
<tr>
<td>MCF-7 (breast epithelium)</td>
<td>FoxA1, AP2γ</td>
<td>oestradiol</td>
<td>ERα</td>
<td>[39]</td>
</tr>
<tr>
<td>U2OS (osteoblast-like cells)</td>
<td>GATA4</td>
<td>oestradiol</td>
<td>ERα</td>
<td>[40]</td>
</tr>
<tr>
<td>3134 (breast epithelium)</td>
<td>AP-1</td>
<td>dexamethasone</td>
<td>GR</td>
<td>[41]</td>
</tr>
<tr>
<td>ESCs</td>
<td>OCT4</td>
<td>TGFβ</td>
<td>Smad2/3</td>
<td>[42]</td>
</tr>
<tr>
<td>myotubes</td>
<td>Myo-D</td>
<td>TGFβ</td>
<td>Smad2/3</td>
<td>[42]</td>
</tr>
<tr>
<td>Pro-B cells</td>
<td>Pu.1</td>
<td>TGFβ</td>
<td>Smad2/3</td>
<td>[42]</td>
</tr>
<tr>
<td>macrophages</td>
<td>Pu.1</td>
<td>LPS</td>
<td>Bcl-6</td>
<td>[31]</td>
</tr>
<tr>
<td>macrophages</td>
<td>Pu.1</td>
<td>LPS</td>
<td>NF-κB</td>
<td>[10]</td>
</tr>
</tbody>
</table>

The enhancer repertoire marked by H3K4me1 evolves during development in response to associated changes in the repertoire of TFs expressed during (and controlling) different stages of differentiation. For example, during the developmental progression from the multipotent haematopoietic progenitor to a committed lymphoid or myeloid cell stage, lineage-restricted enhancer elements were found to be already primed by H3K4me1 prior to differentiation [47]. Upon commitment to the B or myeloid lineages, an increase in H3K4me1 was observed across genomic regions associated with a B-cell or myeloid lineage gene expression programme, respectively. Thus, developmental transitions are associated with dynamic changes in the H3K4me1 repertoire.

More importantly, the mechanism by which the same stimulus is able to activate a distinct enhancers repertoire in different cells has been described. From a conceptual point of view, the general principle is that master regulators are responsible for determining the cell-type-specific effects of the same signalling pathway. In a paradigmatic case, the cell-type-specific effects of TGFβ signalling were found to be determined by the interaction of the TGFβ-activated TFs Smad2/3 with TFs that specify and maintain cell identity [42]. The genomic distribution of Smad3 was found to overlap Oct4 in embryonic stem cells (ESCs), Myod1 in myotubes and Pu.1 in pro-B cells. These cell-type-specific master regulators allow Smad3 binding by establishing open chromatin accessible regions first and by tethering Smad3 to them via protein–protein interactions when the signal is delivered [42].

In the innate immune system, external inputs such as proinflammatory stimuli are expected to have a broad impact on chromatin organization and marking at cis-regulatory elements. In macrophages, a recent genome-wide analysis confirmed that thousands of cis-regulatory regions gain H4 acetylation upon the proinflammatory stimulus lipopolysaccharide (LPS) as expected [48]. These studies also suggested a specific and non-redundant role of the histone deacetylase HDAC3 in controlling acetylation levels at a subset of genomic regions. In this context, it would be of interest to study how distinct stimuli can reorganize the enhancer repertoire of professional innate immune cells to clearly define general principles regulating the dynamic use of the available cis-regulatory information. It is possible that the impact of a broad array of stimuli on such highly specialized cells goes far beyond simple changes in acetylation and results in a reorganization of a subset of enhancers.

### (c) Combinatorial control of transcription factors on cis-regulatory regions

From a mechanistic point of view, pioneer TFs acting as global genomic organizers in a lineage-restricted manner, need additional restricted or non-restricted TFs to bind and activate specific subsets of enhancers. Specific rules controlling functional cooperation in the presence or the absence of cooperative binding remain however to be defined. As discussed above, in macrophages Pu.1 controls the establishment of the enhancer repertoire. However, how it cooperates at a genomic level with other TFs to define or activate specific subsets of enhancers is poorly understood.

In this context, a major breakthrough was recently provided by Amit and co-workers [36] in a study describing genomic occupancy of a large number of TFs in mouse dendritic cells. A high-throughput chromatin immunoprecipitation method (HT-ChIP) was applied to build genome-wide dynamic maps of 25 TFs and four chromatin marks at different time-points following LPS stimulation [36]. Analysis of these dynamics showed that TF binding falls into at least three broad classes. The first class identified TFs with a very pervasive association with almost all regulatory elements in the genome. This group includes the pioneer TF Pu.1 and a very few additional TFs, such as Cebpβ. The broad distribution of these TFs is compatible with their role as chromatin openers that facilitate access of a second group of TFs called ‘primers’, such as Junb, Irf4 and Atf3, that are able to prime for activation regions that are associated with stimulus-dependent gene induction. A third set of TFs (that includes NF-kB and Stat family members) bind dynamically specific set of genes in a stimulus-dependent manner and control gene expression induction (figure 2). Most importantly, the new idea that emerges from this study is that the type of transcriptional response to stimulation is established prior to stimulus delivery. In other words, many TF–DNA interactions are established predominantly at early stages before the treatment with the inflammatory stimulus. Therefore, the potential for proinflammatory genes to be transcriptionally induced is established before stimulation by binding of TF subset to their genomic cis-regulatory regions.
4. Concluding remarks and perspectives

Recent advances in high-throughput sequencing technologies have provided an increasingly sophisticated picture of the mechanisms that regulate inflammatory gene expression programmes in cells of the innate immune system. The data reviewed here indicate that the molecular events occurring during lineage commitment, such as TF binding and changes in chromatin structure, provide competence for subsequent transcriptional induction. However, it is important to conclude by emphasizing that this general framework cannot fully describe the complexity of the inflammatory response and more in general of inducible responses activated by external stimuli. A higher level of data integration, such as that achieved by the ENCODE (Encyclopedia of DNA Elements) consortium project will allow in the future to obtain a more detailed view of coordinated changes in gene expression in response to stimulation [49–52]. In particular, DNAse I footprints coupled with ChIP-seq occupancy maps might be useful to better characterize highly cell-selective and dynamic occupancy patterns and the combinatorial regulation of tissue-specific TF that underlies specific functions.

Additional aspects will have to be included in these models, such as the role of TFs in controlling the three-dimensional network of interactions between distal regulatory elements and target genes, as well as the role of non-coding RNAs, including the enhancer-templated ones [37,53–56]. In this context, topological studies performed in different model system, such as the α- and β-globin gene loci, the immunoglobulin and other antigen receptor gene loci, the imprinted H19–Igf2 locus and the Hox gene clusters, have provided evidence that regulatory DNA sequences can control transcription over a long distance by physically contacting target genes via chromatin looping [57–60]. It will be of interest to define general principles describing how chromatin contacts dynamically change in response to a changing environment, such as the delivery of an inflammatory stimulus.

It is also important to realize that individual genome-wide binding experiments provide a static picture of a dynamic process, where the information of a single TF binding is not enough to predict the regulatory outcome on the target gene. In addition, the current ChIP-seq technologies are not sufficient to determine the true nature of transcription in live cells or at single cell levels. Technical improvements in this technology will allow identifying at greater resolution all TFs binding events in a specific regulatory region and ultimately to understand how they impact on gene expression.

Studies of transcriptional regulation in cells of the innate immune system in GN laboratory are funded by ERC (to G.N.) and Italian Ministry of Health (to S.G.).

References


44. Chen X et al. 2012 Requirement for the histone demethylase Hdac3 for the inflammatory gene


