Biochimica et Biophysica Acta 1849 (2015) 1051–1056

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

Perspective on unraveling the versatility of 'co-repressor' complexes

H. Irem Baymaz [1], Ino D. Karemaker [1], Michiel Vermeulen *

Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 23 June 2015
Accepted 26 June 2015
Available online 28 June 2015

Keywords:
Co-repressor complexes
Histone deacetylation
Gene regulation
Technology integration

ABSTRACT

A multitude of post-translational modifications take place on histones, one of the best studied being acetylation on lysine residues, which is generally associated with gene activation. During the last decades, several so-called co-repressor protein complexes that carry out the reverse process, histone deacetylation, have been identified and characterized, such as the Sin3, N-CoR/SMRT and NuRD complexes. Although a repressive role for these complexes in regulating gene expression is well established, accumulating evidence also points to a role in gene activation. Here, we argue that integration of various state-of-the-art technologies, addressing different aspects of transcriptional regulation, is essential to unravel this apparent biological versatility of 'co-repressor' complexes.

1. Introduction

To facilitate the packaging of DNA into the cell nucleus, four types of histone proteins (H2A, H2B, H3 and H4) together form an octamer around which 147 base pairs of DNA is wrapped. This complex of histones and DNA is called the nucleosome. Both the core domains and the N-termini of histones, the so-called histone tails, contain a large number of amino acid residues that can be post-translationally modified. Because of their proximity to DNA, these modifications can influence gene expression. An important and frequent modification that occurs on histone tails is acetylation, in which an acetyl group ($COCH_3$) is covalently coupled to a lysine residue (K). Already in the 1960s, histone acetylation was proposed to be associated with gene activation [1], an assumption that has been supported by many findings since (e.g. [2]). The acetyl group can neutralise the positive charge of the histone, thus disrupting higher-order chromatin structure by weakening the interaction between the histone octamer and DNA [3]. In general, this leads to enhanced gene transcription *in cis*. In addition, acetylated histones are recognized by bromodomain-containing activator proteins [4]. Many proteins have been identified that have the ability to acetylate histones; these enzymes are called histone acetyltransferases (HATs, reviewed in [3]). Most protein complexes containing HATs are thus known transcriptional activators. In contrast, histone deacetylases (HDACs) remove acetyl groups from histones and these proteins are therefore thought to have a repressive effect on gene expression. Indeed, several complexes containing HDACs were described to induce transcriptional repression, such as the Sin3A/B, N-CoR/SMRT and NuRD complexes [5–11]. Although a repressive role for these complexes in regulating gene expression is well established, accumulating evidence has revealed that co-repressor complexes also localise to actively transcribed genes and are sometimes required for their activation. In this perspective, we discuss some examples where the Sin3A/B, N-CoR/SMRT and NuRD complexes are associated with gene activation (for a more comprehensive review, refer to [12]). We also discuss recent technological advances that need to be further developed and integrated to decipher the molecular mechanisms behind the apparent biological versatility of 'co-repressor' complexes.

2. When co-repressors activate

Early on it was shown that upon inhibition or depletion of HDACs, an approximately equal number of genes are up- and downregulated [13–16]. While these observations may be attributed to indirect effects, subsequent studies proved otherwise [17]. Pioneering mechanistic studies in yeast clearly showed that the Sin3/Rpd3 complex is required for activation of some of its target genes upon osmostress and heat shock in a histone deacetylation-dependent manner [18,19]. The development of new genomics technology also brought about new perspectives on the function of histone acetylation and deacetylation, both in yeast (ChIP-on-chip) and mammals (ChIP-sequencing). The first global localization study for Rpd3, interestingly, showed enrichment both near repressed and active genes [20]. Similarly, in mammals, the genome-wide localization of several HATs and HDACs showed that binding of HATs as well as HDACs positively correlates with gene expression, RNA Polymerase II binding and acetylation levels [21]. The authors proposed that at active genes, the function of HDACs may be to remove acetyl groups during transcription, so that chromatin is 'reset' for the next round of transcription [21]. This suggests that the role of HDACs

* Corresponding author.
E-mail address: M.vermeulen@ncmls.ru.nl (M. Vermeulen).
[1] These authors contributed equally and are listed alphabetically.

http://dx.doi.org/10.1016/j.bbagrm.2015.06.012
1874-9399/

may not be solely repressive, but can in fact also support active gene expression.

Indeed, targeted studies focusing on individual HDAC-containing complexes provided further support for these initial observations. The mammalian orthologues of Sin3, Sin3A and B, were shown to be enriched in the proximity of genes that are activated as myoblasts differentiate into myotubes [22]. A subset of these genes is downregulated upon Sin3A/B depletion, further implying the activating effect of Sin3 proteins on these genes. It is noteworthy that Sin3A/B target genes that are repressed during myoblast differentiation are also enriched for the transcription factor E2F4. Relevant to this observation, in a more recent report by the same group, Sin3A was reported to be enriched at promoters of active genes with high levels of H3K4me3, along with the H3K4me3 interactor ING1 [23]. Another set of genes that are Sin3A targets were also enriched for E2F4. Upon perturbation of the H3K4me3/me1 ratio at the promoter proximal regions, decreased binding of both ING1 and Sin3A was observed in the former set of genes, whereas Sin3A/E2F4 targets were insensitive to the H3K4me3/me1 ratio. These observations classify Sin3 targets depending on different co-factors and/or the chromatin marks that are present *in cis*. Thus, recruitment of Sin3 to chromatin by different co-factors possibly results in different transcriptional outputs.

For the nuclear receptor co-repressors, it was first observed in 2000 that N-CoR was required for transcriptional activation mediated by the retinoic acid receptor-specific ligand (LG550). This ligand failed to activate target genes in N-CoR$^{-/-}$MEFS and exogenous expression of N-CoR restored activation [24]. Similarly, the closely related SMRT co-repressor was observed to mediate full oestrogen-dependent ER-alpha transcriptional activation, albeit in a cell type-specific manner [25]. In addition, a recent ChIP-sequencing study revealed that N-CoR and SMRT colocalize with known activators at Vitamin D3-activated gene enhancers. The genome-wide binding profile of N-CoR showed high correlation with the Vitamin D receptor binding upon Vitamin D3 stimulation, suggesting an activating role for N-CoR [26]. While most of the evidence for an activating role of N-CoR and SMRT is based on correlation, it is clear that gene regulation mediated by these co-repressors is more complex than originally described (reviewed in [27]).

For the NuRD complex, a number of studies have suggested a correlation with active transcription. For example, when the genome-wide localisation of the NuRD subunit MBD3 was studied in MCF-7, MDA-231 and mES cells, it was found that MBD3 preferentially localises to the active part of the genome, i.e. active enhancer regions, active promoters and gene bodies of actively transcribed genes [28,29]. Together with the finding that the NuRD subunit Mi-2β localises mainly to transcriptionally active genes [30], this suggests an activating role for the NuRD complex at these loci. Also in the haematopoietic system, an activating role for the NuRD complex was described. FOG-1, which itself is a co-factor for the haematopoietic factor GATA-1 and is required for regulation of most GATA-1 dependent genes, is known to bind to the NuRD complex and recruit it to FOG-1 target genes [31]. While the NuRD complex is present at target genes of GATA-1 and FOG-1 that are repressed, it is also recruited to some other GATA-1 and FOG-1 targets where it mediates activation. Indeed, it was shown that NuRD occupancy at these genes remained high or even increased upon FOG-1-dependent activation. Additionally, a construct containing a FOG-1-dependent reporter gene could no longer be activated by FOG-1 when the NuRD subunits MTA1-3 were knocked down, indicating that NuRD is directly required for transcriptional activation [31]. By studying mice homozygous for a mutated form of FOG-1 that is unable to interact with the NuRD complex, an activating role for NuRD was suggested for some FOG-1/GATA-1 target genes also *in vivo* [31]. Together, these findings show that the NuRD complex is directly involved in activation of gene expression in multiple instances, indicating that this may be a more general role than previously thought.

3. Data interpretation and future directions

When studying transcriptional regulation by multi-subunit protein complexes, there are many factors that need to be considered that are generally overlooked in genome-wide sequencing based approaches. Below we describe these factors, and discuss several methods and technologies that we think are required, in an integrative manner, to solve the molecular mechanisms underlying the apparent biological versatility of 'co-repressor' complexes (see Fig. 1).

First of all, most genome-wide profiling studies such as ChIP-sequencing require large amounts of cells. Heterogeneity within a certain cell population that is used for the experiment may compromise correct interpretation of the data. In large heterogeneous cell populations, opposing individual signals may be averaged out and therefore escape notice. Also, signals present in a minority of cells may be lost due to overruling signals from the majority. Finally, proteins that appear to co-occupy certain genes in a ChIP-sequencing experiment, may actually bind to the same genes but in a mutually exclusive manner. It is thus important to realise that population effects can strongly affect data output and result in false, generalizing conclusions concerning individual cells.

Asynchronous cell populations cause another layer of heterogeneity. Cells that are in different phases of the cell cycle differ in the genes they express and the epigenetic marks that are present on their chromatin. Thus, pooling asynchronous cells possibly masks some cell cycle specific patterns. A 2003 study using synchronized cell populations showed alternating enrichment of factors driving waves of expression [32]. This led to the model of cyclical regulation of transcription, which was a remarkable leap from the analogy that visualizes transcriptional control as an on/off switch [27]. DNA methylation has also been shown to vary cyclically and thereby regulate gene expression [33,34]. These experiments in synchronized cells thus revealed that cyclical variation may be a more general phenomenon in the regulation of gene expression and that conclusions drawn from the analysis of asynchronous cell populations may not represent all regulatory patterns equally well.

An additional point that needs to be emphasised is that correlation does not necessarily mean causation. When two proteins or epigenetic marks are found to colocalise at a specific locus, this could indeed mean that the presence of one led to recruitment of the other, or that both interacted already before binding at the locus together. However, proteins or marks could also be attracted to the same site independently of each other, for example through binding other proteins. Other scenarios are also imaginable, such as one protein or mark actually causing repulsion of the other, but that this event had not yet taken place at the time of analysis. When examining correlation studies we therefore need to take these different possibilities into account, and appreciate the different conclusions they may propose. Information on the duration of the binding of factors may vastly add to the interpretation of correlation studies. It was first shown in 2000 that regulatory factors are not statically bound at their target regulatory sites upon stimulation, but are rather in a state of constant exchange [35]. Additionally, in the 2003 study mentioned above, the dynamics and co-occurrence of factors on the oestrogen-responsive pS2 gene promoter were identified over a series of time course experiments [32]. Knowing the dynamics of factor binding at target sites provides a more detailed picture of the *in vivo* situation and allows us to decipher the cause within the correlation.

Essential to keep in mind is the structural heterogeneity of the Sin3A/B, N-CoR/SMRT and NuRD complexes. Each complex consists of multiple core subunits, many of which have orthologues, resulting in a large number of possible complex compositions [27,36–38]. Evidence is emerging that similar but distinct core subunits have different biological functions, implying that the exact subunit composition influences the activity of the complex [22,39,40]. In addition, various substoichiometric, often cell-type specific interactors have been identified [38]. These proteins also have the ability to alter the behaviour of the complex, for example by directing the complex to specific targets. Identifying the exact composition of the complexes in different contexts should also shed