An evolving understanding of nuclear receptor coregulator proteins

Christopher J Millard, Peter J Watson, Louise Fairall and John W R Schwabe

Henry Wellcome Laboratories of Structural Biology, Department of Biochemistry, University of Leicester, Leicester LE1 9HN, UK

Correspondence should be addressed to J W R Schwabe
Email john.schwabe@le.ac.uk

Abstract

Nuclear receptors are transcription factors that regulate gene expression through the ligand-controlled recruitment of a diverse group of proteins known as coregulators. Most nuclear receptor coregulators function in large multi-protein complexes that modify chromatin and thereby regulate the transcription of target genes. Structural and functional studies are beginning to reveal how these complexes are assembled bringing together multiple functionalities that mediate: recruitment to specific genomic loci through interaction with transcription factors; recruitment of enzymatic activities that either modify or remodel chromatin and targeting the complexes to their chromatin substrate. These activities are regulated by post-translational modifications, alternative splicing and small signalling molecules. This review focuses on our current understanding of coregulator complexes and aims to highlight the common principles that are beginning to emerge.

Key Words
- nuclear receptors
- coregulator complexes
- coactivators
- corepressors
- chromatin
- transcriptional regulation

Historical perspective

Nuclear receptors are a large family of DNA-binding transcription factors that are recruited to specific DNA sequences in the genome and regulate the expression of genes in the proximity (<1 million base pairs) of these sites. Nuclear receptors were first cloned in 1985. Over the next 10 years our understanding of their molecular mechanisms of action grew steadily and culminated with the determination of the structures of the DNA- and ligand-binding domains (reviewed in Wurtz et al. (1996), Khorasanizadeh & Rastinejad (2001) and Brélivet et al. (2011)). These structures revealed much about how nuclear receptors recognise DNA and their cognate ligands, but did not immediately explain how their recruitment to DNA changes the rate of transcription of their target genes. In pursuit of this understanding, a great many studies sought to identify direct interactions between nuclear receptors and components of the general transcriptional machinery, based on the reasonable presumption that, like prokaryotic transcriptional activators, liganded nuclear receptors would directly recruit RNA polymerase II or associated proteins. These interactions turned out to be elusive, sparking a search for adapter proteins that would make bridging interactions to the general transcriptional machinery. Subsequently, it has become clear that coregulator proteins play a central role in the mediation of transcriptional regulation by nuclear receptors.

The first nuclear receptor coregulators were discovered through efforts to identify proteins that interact with either ligand-bound or unliganded nuclear receptors (Halachmi et al. 1994, Baniahmad et al. 1995, Cavallès et al. 1995). This led to the identification of two classes of
coregulator termed as coactivator proteins, such as SRC1 (Onate et al. 1995), and corepressor proteins, such as NCoR and SMRT (Chen & Evans 1995, Hörllein et al. 1995). We now know that a very large number of proteins participate in the regulatory complexes recruited to nuclear receptors. Indeed more than 350 nuclear coregulators have been identified (http://www.NURSA.org). These proteins are a varied and diverse group that defy straightforward classification. However, it is important to consider some sort of definition for what is meant by coregulator proteins.

The broadest definition is that a coregulator is a protein that participates in a complex that is recruited to the genome by DNA-binding transcription factors, and through this recruitment, the complex regulates (up or down) the rate of transcription of one or more genes. These complexes often have a universal reach throughout the genome with an overarching role in the regulation of transcription (Lonard & O’Malley 2012). It is important to recognise, however, that such a broad definition results in a huge number of proteins being classed as coregulator proteins. Indeed recent proteomic approaches have identified a frighteningly large number (thousands) of proteins as being involved in coregulator complexes (Foulds et al. 2013).

In this anniversary issue, we will attempt to draw out some of the general principles that have emerged concerning the mechanism of action of coregulator complexes, i.e. how they are recruited to specific genomic loci, what enzymatic activities they contain, how the complexes are targeted to chromatin, how coregulator complexes are assembled and finally how coregulators themselves may be regulated. This is not a comprehensive review and there will be an inevitable bias towards how structural biology has given us a better understanding of a small number of these complexes.

**Recruitment to specific genomic loci**

Coregulator complexes do not themselves contain constitutive components that direct sequence-specific recruitment to particular loci in the genome. However, recruitment to nuclear receptors and other sequence-specific transcription factors means that coregulator activity is directed to specific genomic loci leading to specific changes in gene regulation. Each coregulator complex can be recruited to many different transcription factors. Similarly, each transcription factor can often recruit many different coregulator complexes. Furthermore, the transcription factors themselves may be ‘tethered’ to other transcription factors rather than directly bound to DNA themselves. Taken together, this means that the specificity of action of coregulator complexes relies upon a complicated, but specific network of recruitment to various transcriptional regulators.

In the case of nuclear receptors, the presence or absence of a bound ligand determines which nuclear receptor coregulator complexes are recruited. Coregulator complexes that repress transcription are generally recruited to unliganded receptors or receptors bound to inverse agonists. In contrast, coregulator complexes that activate transcription are recruited to receptors bound to agonist ligands. The opposing action of these complexes has led to the widespread use of the terms corepressor and coactivator complexes. However, as we suggest later, it may be better to define complexes based on their mechanism of action rather than the transcriptional outcome.

The molecular mechanisms through which nuclear receptors recruit coregulator complexes are now well-established (reviewed in Watson et al. (2012b)). This is the result of the initial studies that identified critical sequence motifs in coregulator proteins that mediate ligand-dependent interactions with nuclear receptors, followed by crystal structures of the ligand-binding domains of nuclear receptors in complex with these interaction motifs. Remarkably, most of the coactivator complexes recruited to nuclear receptors do so through a conserved motif with the sequence LxxLL (the so-called NR box; Heery et al. 1997). Shortly after the identification of the NR box, an analogous sequence motif (LxxH/I/xxxL/I) was identified in corepressor proteins (Nagy et al. 1999, Hu & Lazar 1999). This motif has been termed as the CoRNR box. Crystal structures have demonstrated that NR box and CoRNR box peptides adopt an α-helical conformation that allows them to bind in the same hydrophobic groove on LBDs and therefore recruitment is mutually exclusive (Nolte et al. 1998, Xu et al. 2002, Wang et al. 2006, Phelan et al. 2010). Which class of coregulator that is recruited to the LBD depends upon the ligand-dependent positioning of the C-terminal helix (often termed helix 12) of the receptor (see Fig. 1).

An important feature of the NR and CoRNR sequence motifs is that they are located within the regions of the coactivator and corepressor proteins that are intrinsically disordered, i.e. do not have an intrinsically fixed structure. The helical structure appears to be formed only on interaction with the nuclear receptor ligand-binding domain. Strikingly, the intrinsic disorder of transcription factor interaction regions is common to many other transcription factor interaction motifs located in these
coregulator proteins (see for example, Ahmad et al. (2003) and Liu et al. (2007)). A consequence of this is that the entropic cost of forming a fixed structure in the bound complex results in a relatively low binding affinity, whilst preserving high specificity, between coregulators and their transcription factor partners. It also allows the coregulator to harbour many transcription factor-binding motifs in intrinsically disordered regions of the protein.

It remains to be seen whether full-length coregulators engage in more extensive contacts with DNA-bound nuclear receptors. The successful determination of crystal structures of full-length nuclear receptors (Chandra et al. 2008, 2013) gives hope that some of these larger complexes may in due course become tractable targets for structure determination.

**Enzymes to modify chromatin**

The original nuclear receptor ‘coregulator concept’ encompassed the idea that if nuclear receptors did not directly interact with the general transcriptional machinery, perhaps there was some protein or protein complex that would act as an intermediary factor. We now know that such intermediary complexes, such as the coregulator complexes are recruited to the genome through DNA-binding transcription factors such as nuclear receptors. (A) Cartoon to illustrate the recruitment of coactivator complex to genes by liganded nuclear receptors, resulting in the acetylation of histone tails and an increase in the rate of transcription. (B) Nuclear receptors recruit corepressor complexes in the absence of ligand, resulting in the deacetylation of histone tails and decrease in the rate of transcription. (C) Crystal structure of PPARγ (light blue) and RXRa (yellow) heterodimer bound to DNA with an NCOA2 peptide (green). Ligands rosiglitazone (for PPARγ and 9-cis retinoic acid (for RXRa are shown as purple sticks). Zinc atoms are shown as grey spheres (PDB code 3DZY (Chandra et al. 2008)). (D) Structure of PPARγ LBD (light blue) with the ligand rosiglitazone (pink) bound to a peptide from the SRC1 coactivator (green) (PDB code 2PRG (Nolte et al. 1998)). (E) Structure of PPARα LBD (light blue) with the antagonist GW6471 (pink) bound to a peptide from the SMRT corepressor (red) (PDB code 1KKQ (Xu et al. 2002)). Helix 12 is highlighted in yellow in both (D and E).

mediator complex, do play an important role in transcriptional regulation. However, the majority of coregulator complexes do not simply act as bridging or assembly factors. Instead they act through bringing enzymatic activities to the locality of the site of their recruitment to the genome. These enzymes either covalently modify chromatin, or result in chromatin reorganisation. The discovery that coregulators act by targeting chromatin represented a highly significant breakthrough in our understanding of the mechanisms of gene regulation. In this review, we will focus mainly on complexes that add, remove and/or read post-translational modifications (PTMs) to nucleosome tails, and thereby influence chromatin accessibility through the further recruitment of other factors to chromatin.

The best-understood modifications include histone acetylation, methylation, phosphorylation and ubiquitylation. In addition to histone modification, DNA itself can be modified through methylation and hydroxymethylation of cytosine. In general, acetylation is associated with transcriptional activation whereas methylation can cause either activation or repression of transcription. Furthermore, histone tail modifications coexist so that particular combinations can bring particular functional outcomes.
For instance, the combined modifications H3K4me2/3+H4K16ac are associated with transcriptionally active homeotic genes, whereas H3K9me3+H3K27me3+CpG 5-MeC is a typical repression mark (Lindroth et al. 2004, Dou et al. 2005). Other combinatorial modifications have been shown to affect mitogen-stimulated transcriptions (H3S10ph+H3K9/14ac) and homeotic gene silencing (H3K27me3+H2AK119ub1) as well as chaperone association, maintenance of heterochromatin and inactivation of the X-chromosome (reviewed in Ruthenburg et al. 2007).

The most intensely studied enzymes that form part of coregulator complexes and are essential in modifying chromatin are the histone acetylases/deacetylases and methylases/demethylases (Bannister & Kouzarides 2011; Fig. 2). These are either termed as epigenetic ‘writers’ and ‘erasers’ depending on whether they add or remove modifications (Gardner et al. 2011, Cosgrove 2012).

Three classes of lysine deacetylases, class I and IIa HDACs, as well as several sirtuin proteins have been shown to act as coregulators of gene transcription. The class I HDACs are zinc-dependent hydrolases that are recruited into at least five different coregulator complexes. In isolation, these enzymes have very limited activity. Once assembled into the complex, the enzymes are activated in part through the binding of inositol phosphate activators (Watson et al. 2012a, Kelly & Cowley 2013, Millard et al. 2013). Like the class I enzymes, class IIa HDACs are recruited to coregulator complexes such as the NCOR complex, but they can also be directly recruited to the genome by repressive transcription factors (Martin et al. 2007). Interestingly, despite their resemblance to class I HDACs, class IIa HDACs exhibit very little HDAC activity (Lahm et al. 2007). Consequently, the mechanism through which they influence transcription remains obscure, yet it is possible that they are activated in some way. Alternatively they may have a different, as yet unidentified, enzymatic activity. Sirtuins have also been shown to influence transcription and can be recruited to coregulator complexes. However, like the class IIa HDACs, this family of NAD-dependent deacetylases does not require obligate assembly into coregulator complexes and can be recruited directly to various transcription factors (Chalkiadaki & Guarente 2012).

Like the class I HDACs, histone acetyl transferases (HATs) are often activated through assembly into larger protein complexes (Yang & Seto 2007). The importance of assembly of HATs into coregulator complexes can be illustrated with Gcn5 (the yeast homologue of PCAF). Purified Gcn5 acetylates free histones but not those assembled into nucleosomes until the coregulator is associated with protein components of the SAGA complex (Grant et al. 1997). In another example, SRC1 is required as an adaptor to enhance the binding of CBP to the estrogen receptor (ER) and therefore direct its HAT activity to the chromatin (Sheppard et al. 2001, Waters et al. 2006).

Despite the lack of detectable primary sequence conservation, structures of the HAT domains from PCAF, p300/CBP and Esa1 show that they contain a structurally conserved central core (Clements et al. 1999, Yan et al. 2000, Liu et al. 2008, Yuan & Marmorstein 2013). Unlike that of the class I HDACs, the catalytic domains of HAT enzymes are often located within larger proteins that contain multiple additional domains. In particular, adjacent bromodomains have been shown to be the important determinants of substrate specificity (Lee & Workman 2007, Zeng et al. 2008). A recent crystal structure of the p300 central region has shown that the HAT domain is tightly associated with the adjacent bromo, PHD and RING finger domains, explaining how the accessory domains direct the HAT activity to correct substrates (Delvecchio et al. 2013).

Methyl groups can be reversibly added to proteins on arginine and lysine residues by peptidylarginine methyltransferases (PRMTs) and lysine methyltransferases respectively. PRMTs require their activity to be directed by accessory proteins, whereas the lysine methyltransferases have some intrinsic targeting ability. For instance, PRMT5 contains only a single domain that must form a stable hetero-octameric complex with the WD40 repeat containing MEPS0 for the direction of methyltransferase to its H4 substrate (Antonymsamy et al. 2012). In contrast, the lysine methyltransferase NSD3 has several domains in addition to the SET domain to direct its own methyltransferase activity; the fifth PHD finger and the CSHCH domain of NSD3 provide the molecular basis for the recognition of unmodified lysine 5 on histone H3 (He et al. 2013). However, while some of the intrinsic targeting activity comes from NSD3 itself, the enzyme associates with the demethylase LSD2 and methyltransferase G9a to establish a pattern of methyl marks on histone tails (Shi et al. 2006).

A family of multi-domain proteins known as lysine demethylases remove the methylation marks placed on coregulators (Cloos et al. 2008). LSD1 and LSD2 are highly related sister enzymes, but assemble into separate protein complexes through the presence or absence of a tower domain. LSD1 is incorporated into the CoREST complex through this ‘stalk-like’ tower domain which forms an
The demethylase activity of its amine oxidase domain is further directed by the adjacent SWIRM1 domain (Stavropoulos et al. 2006). In contrast, LSD2 is recruited through an alternate mechanism that does not require the tower domain, as shown by the crystal structure of LSD2:NPAC:H3 (Fang et al. 2013). NPAC is a putative H3K36me3-binding protein and may provide a mechanism for targeting LSD2 to this histone modification.

ATP-dependent helicases are key components of many coregulator complexes that both remodel chromatin and detach other bound proteins to allow greater access of the transcriptional machinery to the DNA. The ATPase activity of these proteins is directed both through adjacent domains and partner proteins. Each ATPase has
some intrinsic chromatin-targeting ability by virtue of adjacent domains, which include bromodomains (targets acetylated lysines), CHD and PHD domains (targets methylated lysines) and HAND-SANT-SLIDE domains (targets unmodified histone tails) (Clapier & Cairns 2009). Crystal structures, cryo-electron microscopy and site-directed photochemical mapping studies on ISW1a lacking its ATPase domains show how these adjacent domains could act as a "protein ruler" and therefore set the spacing between two adjacent nucleosomes (Yamada et al. 2011). The crystal structure of the eukaryotic CHD1 shows that the ATPase domain interacts with, and is regulated by, the adjacent double chromodomain (Hauk et al. 2010). Further ATPases ISW1 and CHD4 show distinct nucleosome-binding and mobilisation activities that are directed by these associated domains (Brehm 2000). In addition, partner coregulator proteins such as the methyl-CpG binding domain protein MDB2 can direct the helicase activity of CHD4 to methylated DNA (Zhang et al. 1999, Ramirez et al. 2012).

**Chromatin targeting**

A striking feature of many of the enzymes associated with coregulator complexes is the apparent lack of specificity for particular substrates. It has become clear that the specificity of these complexes is rarely determined by the catalytic subunit of the enzyme but is in fact promoted by other domains, present in the coregulator holo-complex, that drive the specificity of interaction. These domains can be located in the same protein as the catalytic subunit or be located in a different subunit of the coregulator complex.

The recruitment of coregulator complexes to the nucleosome is a key step in transcriptional regulation. Many of the multi-domained coregulator proteins contain chromatin-binding modules and a number have been structurally characterised with peptides corresponding to modified and unmodified histone tails. The structure of the bromodomain of GCN5 bound to a peptide corresponding to an acetylated histone H4 peptide shows clearly the mode of recognition and specificity (Owen et al. 2000). The main determinant of the specificity of acetyl lysine recognition is through a number of aromatic residues that form a hydrophobic binding pocket on GCN5 (Fig. 3A). A histidine binding in another hydrophobic pocket confers further specificity.

In the case of the binding of the chromodomain of HP1 to a peptide corresponding to histone H3 methylated at Lys9, the recognition is primarily provided by the methyl group binding in a pocket of three aromatic residues (Fig. 3B; Jacobs & Khorasanizadeh 2002, Nielsen et al. 2002). The major determinant of the specificity appears to be an alanine that is buried in a deep pocket that could not be occupied by any other amino acid.

WD40 domains are common motifs that are present in chromatin-binding proteins, but are distinct from bromodomains and chromodomains. The structure of the WD40 domain polycomb protein EED bound to a peptide corresponding to tri-methylated Lys27 from histone H3 shows the tri-methyl group in an aromatic cage formed by three hydrophobic residues (Fig. 3C; Margueron et al. 2009). The structure of the WD40 domains from RBAP46, and its Drosophila homologue p55 interacting with an H4 peptide, shows an alpha helical structure for the peptide bound to a pocket in the side of the WD40 domain. The structure of RBAP46 suggests that the histone tail is partially unfolded by the binding event (Fig. 3D; Murzina et al. 2008, Song et al. 2008).

More recently structures of chromatin-binding proteins bound to nucleosomes have been obtained and these have provided insight into the mechanism of coregulator complex recruitment. The BAH domain of SIR3 and WD40 domain from the chromatin factor RCC1 both show significant interactions with the histone octamer (Fig. 3E and F). The BAH domain of SIR3 interacts mainly with the tail of H4 that becomes folded upon binding. The conformational change promotes interaction of the H4 tail with the nucleosomal DNA. The BAH domain also interacts with histones H2B and H3 (Armačhe et al. 2011, Wang et al. 2013). The affinity of binding of the BAH domain is increased when the N-terminus is acetylated (Arnaudo et al. 2013, Yang et al. 2013). In addition to extensive contacts to the H2A-H2B histone dimer, the WD40 domain of RCC1 interacts directly with the nucleosomal DNA (Makde et al. 2010).

Many of the coregulator proteins have adjacent domains that contribute to the specificity. For example the methylase JMJD2A has a double tudor domain C-terminal to the catalytic domain. In JMJD2, two tudor domains are interdigitated so that part of each motif forms one domain that binds to histone H3 trimethylated atlysine 4 (Huang et al. 2006). The ADD domain of the heterochromatin protein ATRX has two rigidly oriented domains that interact with two different modifications on the tail of H3. The PHD zinc finger interacts with unmodified H3K4 and the GATA-like zinc finger and its linker with the H3K9Me3 (Eustermann et al. 2011).
Assembling complexes

It has become abundantly clear that most if not all coregulator proteins involved in regulating transcription operate in the form of higher order complexes. As we explore these both structurally and functionally, it emerges that the main reason for this is that multiple functionalities have to be coordinated in order to bring about the required transcriptional outcome. The protein–protein interactions involved in assembling these complexes seem to come in two flavours: high affinity interactions that contribute to obligate and stoichiometric core complexes, and much weaker, transient interactions such as recruitment interactions with transcription factors. Indeed, mass spectrometry has shown that the core of coregulator complexes containing the enzymatic activity is relatively stable, as demonstrated by high resistance to salt treatment, whereas transcription factors preferentially exist in rapid equilibrium (Bantscheff et al. 2011, Joshi et al. 2013).

Electron microscopic analyses of several coregulator complexes, including the SAGA, mediator and PRC2 complexes, have revealed that these complexes, despite containing multiple proteins with regions predicted to be intrinsically disordered, exhibit surprisingly rigid and uniform architectures when fully assembled (Wu et al. 2004, Cai et al. 2012, Ciferri et al. 2012).

The assembly of the obligate core complexes is frequently co-ordinated by hub/platform proteins (Fig. 4) and these mediate interaction with nuclear receptors and other transcription factors (Bulynko & O’Malley 2011). The creation of an active complex through this modular assembly ensures flexibility to complete tasks that range from remodelling of chromatin to modification of individual residues in order to bring about a coordinated response. For example, the deacetylase activity of HDAC3 is directed by hub protein SMRT and two further core proteins TBL1 and GPS2, which together target the HDAC activity to chromatin, resulting in transcriptional...
repression (Oberoi et al. 2011). Once assembled, these coregulator complexes then have the full repertoire of recognition modules and enzymatic factors required to enhance or repress nuclear receptor-mediated transcription.

Dimers and tetramers are a common feature of coregulator core complexes. Many of the enzymatic properties within the complex are duplicated and this allows these multivalent activities to be spread to a larger region of chromatin or to more than a single nucleosome. Tetramerisation of the core of the SMRT repression complex is mediated by the amino-terminal domain of TBL1, which recruits two copies of SMRT and two copies of GPS2, resulting in the formation of a large coregulator complex (Oberoi et al. 2011). The corepressor scaffold protein TUP1 (Schizosaccharomyces pombe) shares a common architecture with TBL1 and both have an amino-terminal tetramerisation domain and a C-terminal WD40. The TUP1 tetramer may oligomerise further to form an extended fibre and therefore accomplish...
long-range transcriptional repression (Matsumura et al. 2012). In a further example, the nucleosome-remodelling enzyme ISWI dimerises when it binds to nucleosomes (Blosser et al. 2009, Racki et al. 2009). Dimerisation improves catalysis and is key to the bidirectional sliding of nucleosomes.

Two-faced coregulators

Historically, coregulators and corepressor complexes have been placed into either the category of coactivator or corepressor, but there are a growing number of scenarios where there is significant disagreement over this classification. Coregulators were originally isolated and defined based on the outcome of single cell-based assays; coactivators are proteins that enhance transcription and corepressors are proteins that repress transcription. There are a number of cases where proteins historically classified as either coactivator or corepressor have been shown to have the opposite effect. For instance, RIP140 appeared to be an inverse coregulator in that it interacts with ligand-bound ER and has the opposite effect. For instance, RIP140 appeared to be an inverse coregulator in that it interacts with ligand-bound ER yet represses transcription (Cavaille`s et al. 1995). Surprisingly, later studies have shown that RIP140 has both coactivator and corepressor properties, which allow it to function as a key and central regulator in development, inflammation and metabolism (reviewed in Nautilia et al. (2013)).

The classical coregulators SRC1 and NCoR/SMRT have also been shown, under certain circumstances, to apparently act in the opposite fashion to their original classification. SRC1, which is most often reported to enhance transcription, can also behave as a corepressor with thyroid hormone receptor (TR) in thyrotropin gene expression (Jeyakumar et al. 1997, Weiss et al. 1999). NCoR/SMRT, which are usually defined as corepressors, can also serve as either TR-dependent activators or repressors, depending on the TR target promoter (Tagami et al. 1997, Berghagen et al. 2002). It is apparent that there are still many questions concerning the underlying mechanisms through which these coregulators function. However, many apparently contradictory findings may be explained through the requirement for rapid acetylation/deacetylation cycling on promoters of nuclear receptor-regulated genes (Burakov et al. 2002, Métivier et al. 2003, Wang et al. 2005).

In some cases, the switch from corepressor to coactivator can be fashioned by the surrounding cellular environment. SMRT can act as a coactivator for full agonist-dependent ER activation in HeLa and MCF-7 cells; however, SMRT reverts to its corepressor role in HepG2 cells, suggesting that the positive effects are cell-type specific (Peterson et al. 2007). In another example, TIP2/GRIP1 is reported as a coactivator for the glucocorticoid receptor and TR, but can act as a repressor when interacting with TR in U2OS cells at the Col3A response element (Rogatsky 2001). This cell-type specificity has been attributed to the availability of gene-specific and cell-dependent factors. These examples highlight that it is important to discuss coregulators as a ‘selective coactivator’ or ‘selective corepressor’ with reference to the cellular environment to clarify the role in that particular cell type.

Regulating the coregulators

Coregulator proteins are, of course, targets of many other regulatory processes. One of the mechanisms for regulating coregulator properties is through the addition or removal of PTMs. These changes to individual surface residues can promote or ablate interactions, altering the composition of protein complexes and allowing the coregulator to rapidly transmit environment signals. In addition, the properties of coregulators can be changed through alternative splicing, proteolytical processing and through the presence of small molecule ligands.

Methylation of several coregulator proteins has been shown to profoundly change their activity. For instance, unmodified RIP140 represses transcription of RARb2, but upon arginine methylation by the methyltransferase PRMT1, this repressive activity is lost (Mostaqul Huq et al. 2006). Arginine methylation has also been shown to affect the activity of further coregulators such as PGC1α (Teyssier et al. 2005) and CBP/p300 (Xu et al. 2001).

Coregulators have also been shown to be regulated by phosphorylation. For example, RIP140 has 11 phosphorylation sites and modification at these sites influences the recruitment of HDACs to the complex (Gupta et al. 2005). In another example, the self-association of SMRT is inhibited upon phosphorylation by the kinase ERK2, causing release from its nuclear receptor partners, and induces a subcellular redistribution of SMRT (Varlakhanova et al. 2011).

Intriguingly, complexes that regulate the acetylation of chromatin have themselves been reported to be regulated by acetylation. For instance, MTA1 is normally classed as a corepressor due to its association with the NuRD repression complex, but can function as a coactivator to stimulate breast cancer amplified sequence 3 (BCAS3) if acetylated at K626. (Gururaj et al. 2006). MTA1 has been suggested to dissociate from its corepressor
complex on acetylation and form an association with Pol II, hence activating transcription.

Ubiquitin and small ubiquitin-related modifier (SUMO) are short polypeptides that are covalently but reversibly attached to proteins that dynamically regulate their localisation and activity (Gill 2004). p300 and CBP have two tandem SUMO sites in their CRD1 domains, and before modification, these proteins facilitate transcriptional activation. However, following SUMOylation, this domain recruits HDAC6 and acts to repress transcription (Girdwood et al. 2003). As well as affecting the activity of coregulators, SUMO modification can retain or direct proteins to different compartments within the cell. SUMOylation of SRC1 was shown to stabilise the association of the progesterone receptor with SRC1 (Chauchereau et al. 2003). Furthermore, SUMOylated SRC1 is retained in the nucleus resulting in enhanced transcription.

PTMs have been shown to act in combination. For example, SRC3 requires phosphorylation followed by the addition of sequential ubiquitin monomers for full activation. However, a sufficiently long ubiquitin chain will signal for the protein for destruction (Wu et al. 2007). This example of dual modification opens up the possibility of a coregulator ‘PTM code’ and suggests that many other marks may be used in combination to influence coregulator function (Lonard & O’Malley 2007). The scope for subtly changing the properties of coregulators through combinations of PTMs has been highlighted through studies on histones (Strahl & Allis 2000, Gardner et al. 2011). It is likely that each individual PTM that has a distinct function may be added to other PTMs to build up this code and massively broaden the functional repertoire of coregulators.

In addition to PTM, proteolytic processing can change the molecular properties of coregulators. A proteolytically cleaved form of SRC1 has altered molecular properties to full length SRC1 and can actively prevent apoptosis (Han et al. 2012). This SRC1 C-terminally truncated protein interacts with proapoptase 8 to prevent activational processing and in this way enhances endometriosis progression.

Differential splicing of coregulators is a further mechanism through which function of coregulator can be regulated. An intriguing example of this is that splice variants change the affinity of NCoR for the nuclear receptors that regulate adipogenesis. The splice variants appear to exert opposing effects on adipocyte differentiation; one splice variant NCoRδ inhibits adipogenesis, while another NCoRα promotes adipogenesis (Goodson et al. 2011). There are at least seven splice variants of NCoR that have been characterised, and these will undoubtedly lead to differential susceptibility to PTMs as well as modulating different interactions with partner proteins (Mottis et al. 2013).

A recent surprise in the field was the discovery that inositol phosphates mediate the interaction between HDAC3 and SMRT (Watson et al. 2012a). Ins(1,4,5,6)P4 binds at the interface between the two proteins and enhances HDAC activity in class I HDAC complexes (Millard et al. 2013). Inositol phosphates have also been shown to stimulate the SWI/SNF complex and therefore seem to have multiple roles in the regulation of chromatin remodelling and transcription (Shen et al. 2003, Steger et al. 2003).

Perspectives

In this review, we have tried to highlight some of the common principles through which coregulator proteins function. The majority of coregulator proteins do not function in isolation, but are assembled into large complexes so as to bring multiple functionalities together. In the broadest terms, these functionalities fall into four classes: i) interactions with transcription factors that recruit the complexes to specific genomic loci; ii) recruitment of enzymatic subunits which when targeted to chromatin either bring about changes in the modification state of the chromatin or lead to chromatin remodelling; iii) interactions with chromatin that are either responsible for recognising the modification state of the chromatin or act as substrate presentation modules facilitating the enzymatic activity of the complex and iv) assembly domains that hold the coregulator complexes together allowing the different functionalities to coexist in one complex molecular machine (Fig. 4).

Despite these common principles, the biological roles and detailed mechanisms of action of the majority of coregulator complexes are yet to be fully understood. One of the biggest questions concerns if (and how) the activities of all these diverse complexes are coordinated? Is there some order of activity? Do these complexes need to be recruited sequentially? Or is it simply a stochastic process? Another important issue concerns the dynamics of recruitment of these complexes to the genome. Several studies have suggested that the residence times of nuclear receptors are very short (review in Stavreva et al. (2012)). How do these residence times relate to the recruitment of coregulator complexes? Does complex recruitment stabilise the transcription factor association with chromatin?
Furthermore, how should we think about the action of these complexes in the complete chromatin landscape? For example, HDACs are classical repressors of transcription, yet they appear to be more abundant on primed promoters and active genes (Wang et al. 2009). Do they perhaps have some role during transcriptional elongation?

Perhaps the most intriguing outstanding issue concerns the emerging evidence that interaction with RNA is an important activity of a number of coregulator complexes. In some cases, it appears that specific long non-coding lncRNAs, such as SRA, play a role in the assembly of coregulator complexes. In other instances, there is evidence to suggest that lncRNAs play a role in recruitment to the genome (Kugel & Goodrich 2012). There is also evidence that there may be RNA-mediated cross-talk with the splicing machinery. The mechanism and biological significance of these interactions remain to be elucidated.

Finally, it is important to note that the driver for understanding the mechanisms and roles of coregulator complexes goes beyond the simple desire for a fundamental understanding, given that there is increasing evidence that compounds which target epigenetic modifiers can be useful in a variety of therapeutic circumstances (Gal-Yam et al. 2008, Helin & Dhanak 2013).

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the review reported.

Funding
This work was supported by the Wellcome Trust (grant number WT085408) and the BBSRC (grant number 10867).


He C, Li F, Zhang J, Wu J & Shi Y 2013 The methyltransferase NSD3 has chromatin-binding motifs, PHD3-C3H4, that are distinct from other NSD (nuclear receptor SET domain) family members in their histone H3 recognition. Journal of Biological Chemistry 288 4692–4703. (doi:10.1074/jbc.M112.426148)


Hu X & Lazar MA 1999 The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. Nature 402 93–96. (doi:10.1038/47069)


Tagami T, Madison LD, Nagaya T & Jameson JL 1997 Nuclear receptor corepressors activate rather than suppress basal transcription of genes
that are negatively regulated by thyroid hormone. *Molecular and Cellular Biology* **17** 2624–2648.


Received in final form 10 October 2013
Accepted 15 October 2013