Cell, Vol. 112, 11–17, January 10, 2003, Copyright ©2003 by Cell Press

Regulating the Regulators: Lysine Modifications Make Their Mark

Review

Richard N. Freiman and Robert Tjian*
Howard Hughes Medical Institute
Department of Molecular and Cell Biology
University of California
Berkeley, California 94720

Decades of research have uncovered much of the molecular machinery responsible for establishing and maintaining proper gene transcription patterns in eukaryotes. Although the composition of this machinery is largely known, mechanisms regulating its activity by covalent modification are just coming into focus. Here, we review several cases of ubiquitination, sumoylation, and acetylation that link specific covalent modification of the transcriptional apparatus to their regulatory function. We propose that potential cascades of modifications serve as molecular rheostats that fine-tune the control of transcription in diverse organisms.

The temporal and spatial control of gene expression is one of the most fundamental processes in biology, and we now realize that it encompasses many layers of complexity and intricate mechanisms. To begin understanding this process, researchers have identified and partly characterized the elaborate molecular apparatus responsible for executing the control of gene expression. This regulatory machinery is a conglomerate of protein factors that function coordinately and in combinatorial fashion to turn specific genes on and off. In multicellular organisms such as humans, it is the selective expression of gene products in individual cell types that leads to the rich diversity of tissues that have evolved to perform highly specialized functions in the body. For example, while brain cells express a unique set of genes important for brain function, liver cells have evolved to express a distinct set of genes important for liver function. To achieve such exquisite control of gene expression, organisms have dedicated a large percentage of their genetic coding capacity to gene products that help determine the spatial and temporal transcription of specific genes.

The molecular machinery responsible for controlling transcription by RNA polymerase II (RNA pol II) is considerably more complex than anyone had anticipated. Over twenty years of transcription biochemistry and genetics has identified a battery of proteins (>80) that aid or abate RNA pol II access to specific regulatory DNA sites, called promoters, scattered throughout the genome where RNA synthesis initiates (see Figure 1A). The first of these regulatory factors to be discovered were the sequence-specific DNA binding transcriptional activators. We now estimate that roughly 5%–10% of the genes in the human genome encode this vast family of transcription factors. Next to be characterized were the general transcription factors, a set of highly conserved proteins responsible for escorting RNA pol II to the promoter. The most recent regulators to be characterized are the coactivators, corepressors, and chromatin remodeling complexes. While the subunit composition and biochemical function of many of these complexes have been well characterized, the precise mechanisms by which these different classes of factors coordinately regulate transcription remain to be elucidated. To uncover some of these mechanisms, a number of recent studies have tracked down the effects of specific covalent modifications of transcription factors in the initiation and elongation phases of the RNA pol II transcription cycle. The critical role of phosphorylation in the control of gene expression has been amply documented (for review see Brivanlou and Darnell, 2002). Here, we will limit our review to other types of modifications such as ubiquitination, sumoylation, and acetylation, which remarkably all target lysine residues and have recently been identified as important modulators of transcription.

While multiple covalent modifications of histone tails have been well characterized and shown to play a global role in gene expression (for review see Berger, 2002; Jenuwein and Allis, 2001), we postulate that modification of nonhistone regulatory proteins (i.e., transcription factors) will play an equally important and perhaps more specific role in directly modulating transcription. The few case studies discussed here will highlight the functional significance of various distinct covalent modifications in the process of transcriptional activation. While the consequences of individual modifications have begun to be addressed, the greater challenge for the future will be to understand how a sequential and possibly a combinatorial network of multiple modifications can regulate gene expression either in a synergistic or antagonistic fashion. Moreover, working out how subtle changes of the transcriptional machinery can vastly alter activation and repression in the context of the large battery of transcriptional initiation factors will be critical to understanding how elaborate gene expression patterns in metazoan organisms are orchestrated.

The Role of Ubiquitin in Regulating Transcription

Ubiquitin, a small, 76 amino acid polypeptide, is covalently linked to lysine residues of target proteins via an enzymatic cascade invoving E1 activating, E2 conjugating, and E3 ligase enzymes (Hershko and Ciechanover, 1998). At first glance, it appeared that transcription factors tagged with ubiquitin would simply be subjected to the classic 26S proteasome-mediated degradation (see Figure 2A). Thus, it was postulated that regulated degradation of ubiquitinated transcription factors would be one way to dictate when a given transcriptional event is activated or deactivated. For example, the processing and nuclear localization of the transcritpion factor NF-κB is controlled by signaling cascades that trigger multiple ubiquitin-coupled and proteasome-dependent degradation events (Hershko and Ciechanover, 1998). In fact, RNA pol II itself is ubqitinated upon DNA damage, and this modification is thought to regulate the proteasomal-

*Correspondence: jmlim@uclink4.berkeley.edu

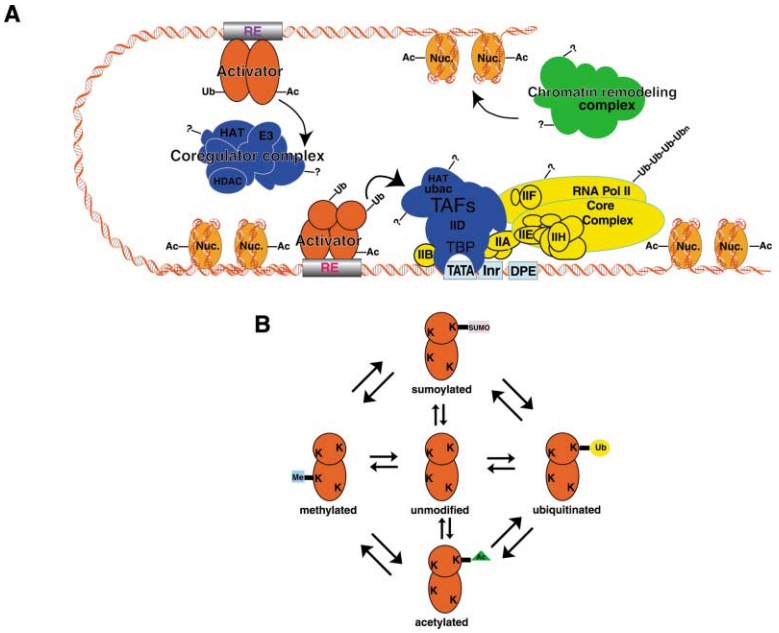

Figure 1. Regulating the Transcriptional Initiation Machinery by Lysine Modification

(A) Potential modification targets of transcriptional regulatory proteins. Depicted are several classes of regulatory factors that cooperate to initiate mRNA synthesis at a typical promoter. These include chromatin remodeling complexes (green), coregulatory complexes including coactivators and corepressors (blue), sequence-specific transcriptional activators (red), and the basal machinery including RNA pol. II (yellow). The covalent linkage of ubiquitin and acetyl groups to potential transcription factors is outlined and putative modifications yet to be identified are indicated by question marks. Important protein-protein interactions that may be affected by covalent modifications are indicated by arrows.

(B) Lysine residues in transcription factors are targets of different covalent modifications. Depicted is a potential network of reversible and dynamic lysine modifications of transcription factors. Lysine residues (labeled K) can serve as the amino acid targets of multiple covalent modifications including sumoylation (SUMO), ubiquitination (Ub), acetylation (Ac), and methylation (Me). An individual lysine residue can only be conjugated by a single modification at a time; however, multiple lysine residues of a protein may be modified simultaneously. Therefore, lysine residues may undergo sequential or cascades of covalent modifications, where modification of an individual residue may influence the modification of a neighboring residue. Abbreviations: Nuc., nucleosome; RE, response element; Ac, acetyl; Me, methyl; Ub, ubiquitin; E3, E3 ligase; HAT, histone acetyltransferase; HDAC, histone deacetylase; UBAC, ubiquitin activating/conjugating activity.

mediated degradation of RNA polymerase II during transcription-coupled repair (Lee et al., 2002; Woudstra et al., 2002). Thus, proteasomal-dependent degradation of the transcriptional machinery is clearly important for numerous mechanisms of transcriptional regulation. However, more recent experiments reveal that under certain circumstances ubiquitination of transcription factors, independent of proteolysis, is also required for the activation function of some transcription factors, suggesting a more complex interplay between regulators.

For instance, transcription of the *MET* genes, which encode enzymes for the biosynthesis of the the sulfur-containing amino acids methionine and cysteine, is coordinately controlled by levels of S-adenosylmethionine (called SAM or AdoMet) in yeast cells. When levels of SAM are low, the transcriptional activator Met4 activates transcription of the genes responsible for methionine biosynthesis. When intracellular levels of SAM rise, these genes are switched off via ubiquitin-dependent inactivation of Met4 by the E3 ligase SCF^{Met30}. Two seemingly contradictory observations suggested that Met4 regulation by ubiquitin is likely to be complex. Rouillon et al. (2000) described the inactivation of Met4 by SCF^{Met30}-dependent ubiqutination followed by degradation of Met4 by the 26S proteasome (Rouillon et al., 2000). In contrast, Kaiser et al. (2000) demonstrated that Met4 inactivation by ubiquitin did not involve proteolysis of Met4, but rather, ubiquitin directly modulates the transcriptional activation function of Met4 (Kaiser et al., 2000). To reconcile these differences, Kuras et al. (2002) recently reported that Met4 is most likely regulated by both degradation-dependent and -independent mechanisms, depending on growth conditions of the cells (see Figure 2B). In minimal media supplemented with methionine, Met4 is ubiquitinated and subsequently destroyed by the proteasome. In contrast, in rich media supplemented with methionine, ubiquitinated Met4 is stable and functional at a distinct set of genes, called the *SAM* genes, but fails to associate properly with the *MET* genes (Kuras et al., 2002). The finding that posttranslational modification of Met4 by ubiquitin controls selective activation of one set of Met4-responsive genes and not another is remarkable and suggests that cells have evolved elaborate mechanisms to coordinately control gene expression but, at the same time, discriminate between different pathways by subtle mechanisms we have only begun to appreciate.

In addition to ubiquitin, components of the proteasome itself have also been implicated in the control of gene expression. The 26S proteasome is composed of a 20S proteolytic barrel and a 19S regulatory sub complex. Johnston and colleagues had previously shown that the 19S regulatory complex is involved in transcriptional elongation by RNA pol II (Ferdous et al., 2001). To further investigate the role of the 19S particle in transcription, Gonzalez et al. (2002) have recently reported that multiple components of the 19S regulatory complex can be recruited to the *GAL1* gene when cells are grown in the presence of galactose. Strikingly, the recruitment of the 19S complex was shown to be independent of the 20S proteasome, suggesting that the 19S activation of the *GAL1* gene is unlikely to only involve proteolysis. The presence of the 19S sub complex at the *GAL1-10* promoter and *GAL1* gene supports its potential role as a direct activator of transcription. However, precise mechanisms of 19S recruitment to promoters potentially involving ubiquitnated activators and the function of this regulatory sub complex in transcriptional initiation and/or elongation remain open areas for future investigation.

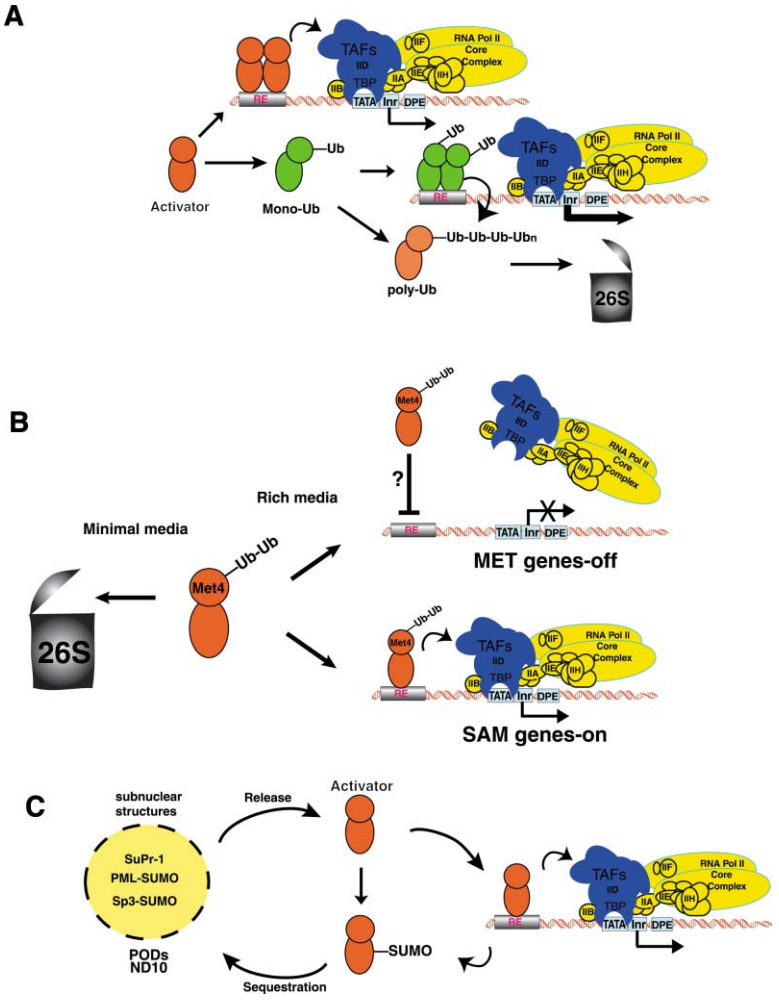

Figure 2. Several Alternative Models Depicting Mechanisms of Transcriptional Regulation by Covalent Modification

(A) Ubiquitination of transcriptional activator proteins. Sequence-specific DNA binding activators can bind to and activate promoters to a certain level. In an undefined manner, monoubiquitination (mono-Ub) is proposed to stimulate the transcriptional potency of activator proteins. Ubiquitin moieties can also be added as long homopolymeric chains, and polyubiquitinated (poly-Ub) transcription factors are likely targeted for proteolytic digestion by the 26S proteasome (depicted as a trash can). Accordingly, ubiquitination can potentially positively and negatively regulate the function of a transcription factor and may be used to precisely regulate the expression of given regulatory pathway in a timely fashion.

(B) Multiple consequences of Met4 regulation by ubiquitin. Met4 activates expression of many genes, including a set of genes encoding enzymes for sulfur amino acid biosynthesis (*MET* genes) and a distinct set of genes required for the production of S-adenosylmethionine (*SAM* genes). The fate of ubiquitinated Met4 is remarkably dependent on growth conditions. In minimal media supplemented with methionine, ubiquitinated Met4 is targeted for degradation by the 26S proteasome. In contrast, in rich media supplemented with methionine, ubiquitinated Met4 is stable. Strikingly, in rich media, ubiquitinated Met4 is differentially recruited to and activates the *SAM* gene promoters and is prevented from association *MET* gene promoters. The mechanism that prevents association of ubiquitinated Met4 with the *MET* genes under these conditions in not known and is indicated by a question mark.

(C) A mechanism of sumoylation-dependent activator regulation. A number of transcription factors are regulated by SUMO modification followed by sequestration in subnuclear structures called PODs or ND10 bodies. When sequestered in such subnuclear compartments, transcription factors are unable to perform their regulatory functions. Subsequent cleavage of SUMO off a given sequestered protein by a SUMO-specific protease allows release of the transcription factor from the PODs, whereby they can resume their regulatory function at the promoter. This proposed cycle of SUMO modification of transcription factors provides a molecular switch by which the cell can respond to alternative stimuli.

In addition to these studies, a number of groups have probed the potential role of ubiquitination on the activity of transactivation domains derived from mammalian viral and cellular activators. Several investigators observed that often the most potent transcriptional activators expressed in cells were also the most difficult to detect by immunoblotting. Looking at this phenomenon in greater detail, Molinari et al. (1999) observed a positive correlation between the potency of activation domains and proteasome-mediated degradation. In addition, Salghetti et al. (2000) found a striking overlap in several transcription factors between the amino acid sequences involved in transcriptional activation and degradation. Together, these studies suggested that strong transcriptional activators may be switched off by proteasome-mediated degradation. But again, one must ask, was simple proteolysis all that was going on? The answer appears to be no. Looking at activation by the synthetic activator LexA-VP16 in yeast cells, Tansey and colleagues observed that direct fusion of ubiquitin to the potent activator is able to bypass the need for a specific ubiquitin ligase, implicating a direct role for ubiquitnation in transcriptional activation (Salghetti et al., 2001). This study suggests that monoubiquitination may enhance the activity of transcription factors and that polyubiquitination may subsequently signal destruction by the 26S proteasome (see Figure 2A). In addition, many deubiquitinating enzymes function to remove ubiquitin from target proteins. Since ubiquitination is a reversible modification, there may be distinct cues that signal the addition and removal of ubiquitin to and from a particular transcription factor. Subtle transition between the non-, mono-, and polyubiquitinated states may act as an important switch to turn on and off a particular trancription factor. Perhaps under certain conditions, E3 processivity factors limit the extent to which ubiquitin is attached to a target, thereby regulating the transition from mono- to polyubiquitination. Future studies of ubiquitination intermediates of natural endogenous activators will be crucial to substantiate the importance of this proposed mechanism in vivo. In addition, uncovering the physiological signals responsible for the activation-coupled

degradation of transcriptional activators will help determine how such pathways are regulated.

Not only is the transcriptional machinery targeted by ubiquitination, a number of recent reports indicate that contained within the regulatory apparatus itself are enzymes that can catalyze the addition of ubiquitin to various target proteins. Brower et al. (2002) recently reported that the MED8 (also called ARC32) subunit of a mammalian cofactor complex is able to reconstitute ubiquitin ligase activity in vitro. Moreover, the MED8-containing complex from rat liver nuclear extracts copurified with an ubiquitin ligase activity is suggestive of an endogenous association between a mammalian transcriptional cofactor and an ubiquitin ligase (Brower et al., 2002). In a related scenario, the largest subunit of the *Drosophila* TFIID complex, dTAF1 (formerly $dTAF_{II}250$), has been shown to harbor E1 and E2 ubiquitin-activating/conjugating (UBAC) activities (Pham and Sauer, 2000). The dTAF1 UBAC activity is observed to be important for dorsal-mediated transactivation and proposed to catalyze monoubiquitination of histone H1. In an independent study, Albert et al. (2002) recently identified an E3 ligase activity within the CNOT4 subunit of the CCR4-NOT transcriptional repressor complex. The function of the CNOT4 E3 ligase is proposed to catalyze ubiquitination of select components of the transcription initiation machinery. Taken together, these studies suggest that the transcriptional machinery itself harbors ubiquitin-modifying enzymes that may regulate the activity of specific subunits of the regulatory apparatus (see Figure 1A). However, deciphering the in vivo targets of such ubiquitin-modifying activities and elucidating their effects on transcriptional activation and repression remains a formidable and critical task.

SUMO Enters the Nucleus

Besides ubiquitination, a number of transcription factors have been documented to be modified by the small ubiquitin-related modifier (SUMO). Sumoylation, the process of conjugating SUMO, a 101 amino acid polypeptide, to target proteins is highly similar to that of ubiquitination (Kim et al., 2002). Both ubiquitin and SUMO are covalently linked to lysine residues of target proteins and adopt similar tertiary structures. In contrast to their structural similarities, the biological consequences of sumoylation versus ubiquitnation are apparently markedly distinct. Proteins tagged with SUMO do not apparently undergo proteasomal-dependent degradation. Instead, sumoylation is often involved in directing the subcellular localization and stabilization of transcription factors. For example, a number of sumoylated transcription factors, including the promyelocytic leukemia gene product PML, have been detected in subnuclear structures variously called PML oncogenic domains (PODs), nuclear bodies, or ND10 structures. The disruption of nuclear PODs observed in acute promyelocytic leukemia suggests that PODs perform a critical function in protecting against certain forms of leukemia. Since PODs contain a number of sumoylated transcription factors, it is thought that sumoylation targets specific transcription factors into specialized subnuclear bodies, thereby regulating their activity and function in transcription (see Figure 2C).

A number of transcription factors regulated by sumoylation have recently been described. For example, the transcription factor LEF1, a downstream effector of the Wnt signaling cascade, has been reported to undergo regulation by sequestration into nuclear bodies in similar fashion to PML. Sachdev et al. (2001) reported that LEF-1 associates with the SUMO E3 ligase PIASy. This association is thought to stimulate sumoylation of LEF-1 and inhibit LEF1-mediated transactivation by sequestration into nuclear bodies (Sachdev et al., 2001). The physiological signals that might trigger release of LEF1 from these nuclear bodies remain to be elucidated, but presumably they will be triggered during the Wnt signaling cascade when LEF1 is activated. Another transcription factor regulated by sumoylation is Sp3, one of several members of the Sp1 family of GC box binding transcriptional regulators. Two recent studies report that Sp3 activity is modulated by sumoylation of a specific lysine residue in an inhibitory domain of the protein (Ross et al., 2002; Sapetschnig et al., 2002). Sapetschnig et al. (2002) found that sumoylation of Sp3 by the SUMO E3 ligase PIAS1, one of several related SUMO E3 ligases, inhibits the transactivation potential of Sp3. In addition to detecting sumoylation-dependent inhibition of Sp3 activity, Ross et al. (2002) detect Sp3 at the nuclear periphery and within nuclear dots in a sumoylation-dependent manner. Together, these studies suggest that sumoylation-associated sequestration of transcription factors in subnuclear bodies may modulate the activity of transcriptional activators and that some as yet unknown signaling event releases these factors to perform their functions.

Are nuclear PODs merely storage compartments of nuclear proteins, or is the release of transcription factors from such sites regulated by physiological cues that are important for transcription? Recent characterization of the mammalian SUMO-1 protease SuPr-1 suggests that regulation of gene expression and nuclear POD formation is functionally linked. Best et al. (2002) report that SuPr-1 is localized to nuclear PODs, associates with and hydrolyzes SUMO-1-modified forms of PML, resulting in the redistribution of POD-localized transcription factors in the nucleus. The positive effect of SuPr-1 on activation is shown to be dependent on PML, as it is lost in PML-deficient mouse embryonic fibroblasts. These data suggest that assembly of transcription factors in nuclear PODs is critical for their subsequent function as activators (see Figure 2C). Remarkably, SuPr-1 was also shown to activate Sp3-mediated transactivation by catalyzing the removal of SUMO-1 from Sp3 (Ross et al., 2002). Expression of SuPr-1 redistributed Sp3 from concentrated nuclear dots and the nuclear periphery to diffuse nuclear staining, allowing Sp3 to efficiently activate transcription. Characterizing the biological context in which these activities are modulated in health and disease will likely uncover novel modes of gene regulation as well as the molecular basis of certain diseases such as leukemia.

Studies by Courey and colleagues on the SUMO (also called Smt3) modification of the *Drosophila* transcription factor Dorsal has revealed that, like ubiquitnation, sumoylation may modulate multiple activities of individual transcription factors. Accumulation of Dorsal, a fly homolog of vertebrate NF-κB, is negatively regulated by

a cytoplasmic binding partner Cactus, in analogous fashion to IκB-mediated inhibition of NF-κB. The covalent modification of Dorsal by SUMO was initially found to relieve Cactus-mediated repression of Dorsal, allowing nuclear translocation and subsequent activation by Dorsal (Bhaskar et al., 2000). However, more recent findings suggest that in addition to relieving Cactus-mediated repression, sumoylation may also enhance the activation function of nuclear-localized Dorsal, possibly by disrupting its association with an unknown nuclear repressor (Bhaskar et al., 2002). Furthermore, Bhaskar et al. (2002) have also implicated the specific sumoylation of Dorsal to its function in regulating the innate immune response in *Drosophila*. Such regulation by covalent modification of transcription factors highlights the potential role of modification dynamics in determining the biological output of the regulatory machinery. Physiologic signals that increase or decrease the extent of sumoylation may thus have a significant influence on the activity of select transcription factors. Future characterization of the functional significance of such regulatory marks will undoubtedly yield novel modes of transcriptional regulation.

Acetylation of Nonhistone Targets

In addition to ubiqitination and sumoylation, lysine residues of regulatory proteins can be covalently modified by acetylation. The acetylation state of lysine residues on histone tails has been well documented (for review see Berger, 2002; Jenuwein and Allis, 2001). In contrast, much less is known about the nonhistone substrates of acetyltransferases and deacetylases, the enzymes that catalyze the addition and removal of acetyl groups, respectively. Numerous histone acetyltransferases (HATs) and histone deacetylases (HDACs) are intimately associated with various components of the core transcriptional machinery. For example, the TAF1 subunit of TFIID has been shown to contain HAT activity, suggesting that it may function to acetylate transcriptional components in addition to histones (Mizzen et al., 1996). Accordingly, the addition and removal of acetyl groups to and from transcription factors will play as important a role in transcriptional regulation as histone acetylation and deacetylation. For example, the HAT-containing coactivator p300 has been documented to acetylate the p53 tumor suppressor protein thereby stimulating its function as an activator (Barlev et al., 2001; Gu and Roeder, 1997). Thanos and colleagues have also demonstrated that the acetylation state of HMGI(Y) is critical for regulating the assembly and disassembly of an enhancesome at the IFN-β gene, further underscoring the importance of nonhistone acetylation in transcriptional regulation (Munshi et al., 2001). Recently, Bereshchenko et al. (2002) reported the ability of p300 to acetylate the transcriptional repressor BCL6. In this case, acetylation inactivates the ability of BCL6 to recruit HDACs and function as a potent repressor of transcription (Bereshchenko et al., 2002). Together, these studies illuminate the ability of HATs to either stimulate or inactivate the function of nonhistone regulatory proteins.

Transcriptional activation by the herpes simplex virus (HSV) activator VP16 includes recruitment of HAT-containing coactivator complexes to viral immediate-early gene promoters as measured by chromatin immunoprecipitation during early HSV infection. (F.J. Herrera and S.J. Triezenberg, personal communication). However, the HSV genome is apparently unadorned by nucleosomes, suggesting that the function of HAT recruitment in the case of VP16 may be to acetylate nonhistone components of the transcriptional initiation machinery. Unraveling the molecular mechanisms of HSV immediate-early gene transcription should reveal the nonhistone substrates of these HAT activities that may be critical for VP16-dependent transactivation. Such mechanisms will likely be relevant to the regulation of cellular gene transcription as well. Thus, we anticipate that the acetylation state of promoter recognition proteins as well as other transcription factors will emerge as a rich area of important future investigations into gene regulatory pathways.

Although we have not included a discussion here, nonhistone methylation is also likely to play a role in transcriptional regulation. In fact, the regulatory activities of CBP/p300 were recently demonstrated to be modulated by arginine methylation (Xu et al., 2001). It is thus likely that histone methyltransferases, like acetyltransferases, will also be found to target nonhistone regulatory proteins and modulate their function accordingly.

Gridlock at Lysine Residues

Since ubiquitination, sumoylation, and acetylation can all occur on lysine residues, transcription factors can potentially undergo a cascade of modifications that modulate their function. This obviously complicates the contribution of each individual modification, and it will be important to sort out the order and dynamics of multiple modification events on endogenous regulatory proteins (see Figure 1B). For example, the major site of Sp3 sumoylation is identical to the major site of acetylation, and both of these modifications have been demonstrated to modulate activation (Braun et al., 2001). Presumably, in such cases one modification may preclude the other or alternatively may be responsible for enhancement of a second modification at a nearby residue. For example, sumoylation of a lysine residue in IκBα has been demonstrated to block ubiquitnation at this identical residue, protecting IκBα from proteasomal-mediated degradation (Desterro et al., 1998). In addition, acetylation of the Smad7 transcriptional regulator has recently been shown to protect Smad7 from ubiquitin-mediated degradation, suggesting a competition between ubiquitination and acetylation at critical lysine residues (Gronroos et al., 2002). Likewise, it has been proposed that sumoylation of a specific lysine residue may protect a given protein from ubiquitination and subsequent proteasomal-mediated proteolysis. It is not hard to envision that these lysine residues therefore serve as critical molecular switches that can respond to different signals in highly specific ways. In addition, since most proteins contain many lysine residues, transcription factors may undergo multiple modifications simultaneously or in sequential order, pointing to the possibility of generating complex networks of regulatory events (see Figure 1B). Sorting out such a molecular switch board, both biochemically and genetically, poses

a formidable but necessary task if we are to understand how tissue-specific and gene-selective transcriptional regulation is achieved.

While the modifications described here all target lysine residues, they are inherently quite different in the manner they alter target proteins. While acetylation entails the addition of a relatively small, uncharged acetyl group, addition of ubiquitin and SUMO involves attaching relatively large polypeptides that can significantly add to the mass of target proteins. Not surprisingly, the functional outcomes resulting from these diverse covalent changes can be very different. One can imagine that such biochemically distinct changes mediate important differences in the function of modified proteins. For example, some types of alterations may induce conformational changes while other modifications may trigger charge surface changes. The net outcomes, however, would be to disrupt and reorganize important protein-protein and protein-DNA contacts. Documenting the biophysical changes in modified proteins should help illuminate the function of these diverse modifications.

Concluding Remarks

The recent completion of several animal genome sequences has revealed that the number of expressed genes is considerably lower than expected. Consequently, the vast differences in cell types, signal transduction pathways, and complex behaviors characteristic of different species cannot be easily explained by increased gene number (i.e., worms, 19K and humans, 30K) alone but, rather, how a relatively limited number of genes (roughly 10–30K) are differentially expressed and utilized. In other words, the dramatic phenotypic differences between a worm and a mammal can at least partially be rationalized by differences in the complexity of the regulatory code and not merely gene content. The discovery of multiple covalent modifications of the regulatory apparatus discussed here suggests that organisms have evolved various mechanisms to maximize the usage of a relatively limited number of genes and transcription factors. By utilizing multiple distinct mechanisms to modify and control the transcriptional machinery, organisms have evolved much greater potential for directing diverse expression profiles by a finite number of transcription factors. Taking advantage of multiple covalent modifications of transcription factors organisms have effectively gained the ability to utilize the same regulatory factor in different ways and thus expand their range of gene expression patterns. Regulation by modification not only enhances the functional potential of each individual transcription factor but also provides an effective means of greatly amplifying the functional plasticity of the transcriptional machinery required for combinatorial diversity. This quantum increase in the repertoire of regulatory events ultimately provides the rich tapestry of molecular interactions necessary to direct the diverse arrays of gene expression programs that define complex organisms.

While many components of the transcriptional machinery are conserved through evolution, we suspect that some modification networks may be specific to individual organisms, resulting in different gene expression outcomes depending on the species. For example, the transcription factor TFIID is largely conserved from yeast to humans; however, the diverse programs of gene expression regulated by this multiprotein coactivator complex in unicellular and multicellular organisms have diverged substantially. Therefore, it is possible that covalent modification of transcription factors, like TFIID, may occur in a species-specific manner, thereby allowing these factors to evolve specialized functions related to their evolutionary niche. For example, TFIID derived from animal cells containing cell type-specific subunits may be modified by specific mechanisms specialized to function in certain tissue types. Clearly, transcription is exquisitely regulated in all organisms, and one mechanism utilized to achieve such regulation is covalent modification of the transcriptional machinery. Future studies in diverse organisms and specialized regulatory pathways should further illuminate how transcription factor modification contributes to the elaborate mechanisms of gene regulation.

References

Albert, T.K., Hanzawa, H., Legtenberg, Y.I., de Ruwe, M.J., van den Heuvel, F.A., Collart, M.A., Boelens, R., and Timmers, H.T. (2002). Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. EMBO J. *21*, 355–364.

Barlev, N.A., Liu, L., Chehab, N.H., Mansfield, K., Harris, K.G., Halazonetis, T.D., and Berger, S.L. (2001). Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. Mol. Cell *8*, 1243–1254.

Bereshchenko, O.R., Gu, W., and Dalla-Favera, R. (2002). Acetylation inactivates the transcriptional repressor BCL6. Nat. Genet. *32*, 606–613.

Berger, S.L. (2002). Histone modifications in transcriptional regulation. Curr. Opin. Genet. Dev. *12*, 142–148.

Best, J.L., Ganiatsas, S., Agarwal, S., Changou, A., Salomoni, P., Shirihai, O., Meluh, P.B., Pandolfi, P.P., and Zon, L.I. (2002). SUMO-1 protease-1 regulates gene transcription through PML. Mol. Cell *10*, 843–855.

Bhaskar, V., Valentine, S.A., and Courey, A.J. (2000). A functional interaction between dorsal and components of the Smt3 conjugation machinery. J. Biol. Chem. *275*, 4033–4040.

Bhaskar, V., Smith, M., and Courey, A.J. (2002). Conjugation of Smt3 to dorsal may potentiate the Drosophila immune response. Mol. Cell. Biol. *22*, 492–504.

Braun, H., Koop, R., Ertmer, A., Nacht, S., and Suske, G. (2001). Transcription factor Sp3 is regulated by acetylation. Nucleic Acids Res. *29*, 4994–5000.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. Science *295*, 813–818.

Brower, C.S., Sato, S., Tomomori-Sato, C., Kamura, T., Pause, A., Stearman, R., Klausner, R.D., Malik, S., Lane, W.S., Sorokina, I., et al. (2002). Mammalian mediator subunit mMED8 is an Elongin BC-interacting protein that can assemble with Cul2 and Rbx1 to reconstitute a ubiquitin ligase. Proc. Natl. Acad. Sci. USA *99*, 10353–10358.

Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Mol. Cell *2*, 233–239.

Ferdous, A., Gonzalez, F., Sun, L., Kodadek, T., and Johnston, S.A. (2001). The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. Mol. Cell *7*, 981–991.

Gonzalez, F., Delahodde, A., Kodadek, T., and Johnston, S.A. (2002). Recruitment of a 19S proteasome subcomplex to an activated promoter. Science *296*, 548–550.

Gronroos, E., Hellman, U., Heldin, C.H., and Ericsson, J. (2002).

Control of Smad7 stability by competition between acetylation and ubiquitination. Mol. Cell *10*, 483–493.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595–606.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annu. Rev. Biochem. *67*, 425–479.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Kaiser, P., Flick, K., Wittenberg, C., and Reed, S.I. (2000). Regulation of transcription by ubiquitination without proteolysis: Cdc34/SCF(Met30)-mediated inactivation of the transcription factor Met4. Cell *102*, 303–314.

Kim, K.I., Baek, S.H., and Chung, C.H. (2002). Versatile protein tag, SUMO: its enzymology and biological function. J. Cell. Physiol. *191*, 257–268.

Kuras, L., Rouillon, A., Lee, T., Barbey, R., Tyers, M., and Thomas, D. (2002). Dual regulation of the met4 transcription factor by ubiquitin-dependent degradation and inhibition of promoter recruitment. Mol. Cell *10*, 69–80.

Lee, K.B., Wang, D., Lippard, S.J., and Sharp, P.A. (2002). Transcription-coupled and DNA damage-dependent ubiquitination of RNA polymerase II in vitro. Proc. Natl. Acad. Sci. USA *99*, 4239–4244.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., et al. (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261–1270.

Molinari, E., Gilman, M., and Natesan, S. (1999). Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency in vivo. EMBO J. *18*, 6439–6447.

Munshi, N., Agalioti, T., Lomvardas, S., Merika, M., Chen, G., and Thanos, D. (2001). Coordination of a transcriptional switch by HMGI(Y) acetylation. Science *293*, 1133–1136.

Pham, A.D., and Sauer, F. (2000). Ubiquitin-activating/conjugating activity of TAFII250, a mediator of activation of gene expression in Drosophila. Science *289*, 2357–2360.

Ross, S., Best, J.L., Zon, L.I., and Gill, G. (2002). SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. Mol. Cell *10*, 831–842.

Rouillon, A., Barbey, R., Patton, E.E., Tyers, M., and Thomas, D. (2000). Feedback-regulated degradation of the transcriptional activator Met4 is triggered by the SCF(Met30)complex. EMBO J. *19*, 282–294.

Sachdev, S., Bruhn, L., Sieber, H., Pichler, A., Melchior, F., and Grosschedl, R. (2001). PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. Genes Dev. *15*, 3088–3103.

Salghetti, S.E., Muratani, M., Wijnen, H., Futcher, B., and Tansey, W.P. (2000). Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis. Proc. Natl. Acad. Sci. USA *97*, 3118–3123.

Salghetti, S.E., Caudy, A.A., Chenoweth, J.G., and Tansey, W.P. (2001). Regulation of transcriptional activation domain function by ubiquitin. Science *293*, 1651–1653.

Sapetschnig, A., Rischitor, G., Braun, H., Doll, A., Schergaut, M., Melchior, F., and Suske, G. (2002). Transcription factor Sp3 is silenced through SUMO modification by PIAS1. EMBO J. *21*, 5206–5215.

Woudstra, E.C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J.Q. (2002). A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. Nature *415*, 929–933.

Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B.M., Montminy, M., and Evans, R.M. (2001). A transcriptional switch mediated by cofactor methylation. Science *294*, 2507–2511.