Dose-dependent inhibition of histone deacetylases reprograms gene expression through global remodeling of the enhancer landscape

by

Gilson John Sanchez

B.S., University of California, Riverside, 2004

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

Department of Chemistry and Biochemistry

2016

This thesis entitled:

Dose-dependent inhibition of histone deacetylases reprograms gene expression through global remodeling of the enhancer landscape

Written by Gilson John Sanchez
has been approved for the Department of Chemistry and Biochemistry by:

_____
Xuedong Liu, Ph.D.

_____
Robin D. Dowell, Ph.D.

Date_____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above-mentioned discipline.

Abstract

Sanchez, Gilson John (Ph.D., Biochemistry)

Dose-dependent inhibition of histone deacetylases reprograms gene expression through global remodeling of the enhancer landscape

Thesis directed by Xuedong Liu, Ph.D.

Histone deacetylases (HDACs) regulate gene expression through deacetylation of histones and non-histone proteins. Histone deacetylase inhibitors (HDACIs) are known to alter gene expression by both up- and down-regulation of protein coding genes in normal and cancer cells. However, the mechanisms that are responsible for activation and repression of gene expression remain obscure. To understand the underlying mode of action of HDACIs, I used genome wide chromatin immunoprecipitation sequencing (CHIP-seq) to determine dose-dependent changes in histone acetylation and methylation marks in HCT116 cells treated with increasing concentrations of the natural product largazole, a class I and class IIb selective HDAC inhibitor. Changes in mRNA expression were also measured by RNA-seq under similar conditions. I found that cells exposed to low nanomolar concentrations ($<GI_{50}$) predominantly resulted in the upregulation of gene transcripts whereas mid to high doses ($\geq GI_{50}$) triggered a decrease in mRNA accumulation.

Largazole's effect on transcription is likely associated with its activity on enhancer elements rather than on gene bodies. Low dose largazole exposure elevates the activity of ~1600 poised enhancers characterized by increased H3K9/27ac and H3K4me2 marks in addition to RNA pol II recruitment, suggesting class I HDACs are involved in maintaining the repressive state of poised enhancers. Mid to high dose largazole appears to directly repress transcription of many genes, an effect correlated with increased RNA Pol II pausing at the promoters of most

actively transcribed genes. Under such conditions, I found ~800 putative enhancers become decommissioned with the features of loss of H3K9/27ac, reduction of H3K4me2 and the decrease of RNA Pol II occupancy. A significant number of primed (~47%) and decommissioned enhancers (~22%) display the recognition motif for the AP-1 transcription factor. Therefore, HDAC inhibitors sculpt the enhancer landscapes in a dose-dependent manner to impact gene expression and cytostatic responses.

Collectively, the data presented here challenge traditional assumptions that HDACI-driven chromatin hyperacetylation of promoters and gene regions results in positive stimulation of transcription. This study represents the most comprehensive analysis to date showing largazole dose dependent genome-wide acetylation changes and has helped identify the activation of a large *cis*-regulatory network as a novel output from HDAC inhibition.

Acknowledgements

I would like to thank my advisor Dr. Xuedong Liu, for his support, for encouraging me to branch into new fields of research, and for allowing me to grow as a scientist. I thank my thesis committee members, Dr. Robin Dowell, Dr. Thomas Cech, Dr. James Goodrich, and Dr. Dylan Taatjes for the conversations and guidance over the years.

I am immensely grateful to my family, who deserves special recognition, since they have been fundamental in the path to achieving my goals. In Boulder, I had the fortune to meet my lovely and always supporting partner, Candice, who has given me the most precious gift of my life, our daughter Ofelia, thank you. To my parents, Gilberto and Josefina for their unconditional love and for the countless days of hard work, thank you. My sisters, Connie and Nancy, who have thought me to slow down and savor precious moments outside the lab, and to my brothers in–law, Humberto and Efrain, I thank you.

I want to especially thank Dr. Natalie Ahn, who first made my transition to Boulder the friendliest event, later patiently helped me thrive through my second year, and ultimately thought me that it is possible to simultaneously excel in– as well as outside the lab.

I feel extremely fortunate to have conducted my Ph.D. at University of Colorado, Boulder, as I have had the fortune to interact with tremendous people that also happened to be exceptional scientists. Doug Chapnick, Eric Bunker, Conggang Zhang, Adrian Ramirez, Phillip Richmond, Joey Azofeifa, Dan Medeiros, Tyler Square, and Aaron Garnett have all share my frustrations as well as those fleeting eureka moments.

Finally, I want to thank Dr. Eduard Platzer at University of California, Riverside for introducing me to the addicting world of scientific research and specially thank Dr. Sandy Sharp at CSU, Los Angeles for helping shape the opportunities that allowed me to arrive to this point.

Thesis Contents

List of Tables

Chapter 1

Chapter 2

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 1: Introduction – background and significance

1.1 Genome organization and accessibility

In eukaryotic cells, genomic DNA is tightly associated with histone proteins in repetitive units known as nucleosomes. This structure is the basic unit of chromatin and it is composed of 147 base pairs of DNA wrapped around a histone octamer made of two copies of each of the core histones (H2A, H2B, H3 and H4) (Luger et al., 1997). Histone H1 serves as a linker protein and is bound to the outside of the DNA-core complex, forming a full nucleosome (Oudet et al., 1975). In tightly packaged chromatin, nucleosomes are found every 200bp, allowing for highly compacted DNA within the nucleus of the cell. The wrapping of DNA around histones and the formation of higher order chromatin structures permits mammalian cells to pack over 3 billion base pairs of genetic information into a nucleus of approximate 8 microns in diameter (Figure 1.1). Although the large amount of chromatin allows for higher complexity, it also introduces an intrinsic problem: the compacted DNA must still interact with a plethora of nuclear factors that direct critical cellular processes such as DNA replication, recombination, transcription, and repair (Taft et al., 2007). This makes the establishment and regulation of chromatin structure a fundamental problem in eukaryotes, as the DNA molecule must be both packaged and readily accessible.

The mechanism by which chromatin allows site-specific access to regulatory proteins is mostly controlled by chromatin remodeling proteins (Bannister and Kouzarides, 2011). The most studied class of chromatin remodelers are histone modifying enzymes which primarily target histone tails for the addition and removal of functional groups. Histone tails are 20 to 40 amino acids in length and readily interact with modifying and regulatory proteins.

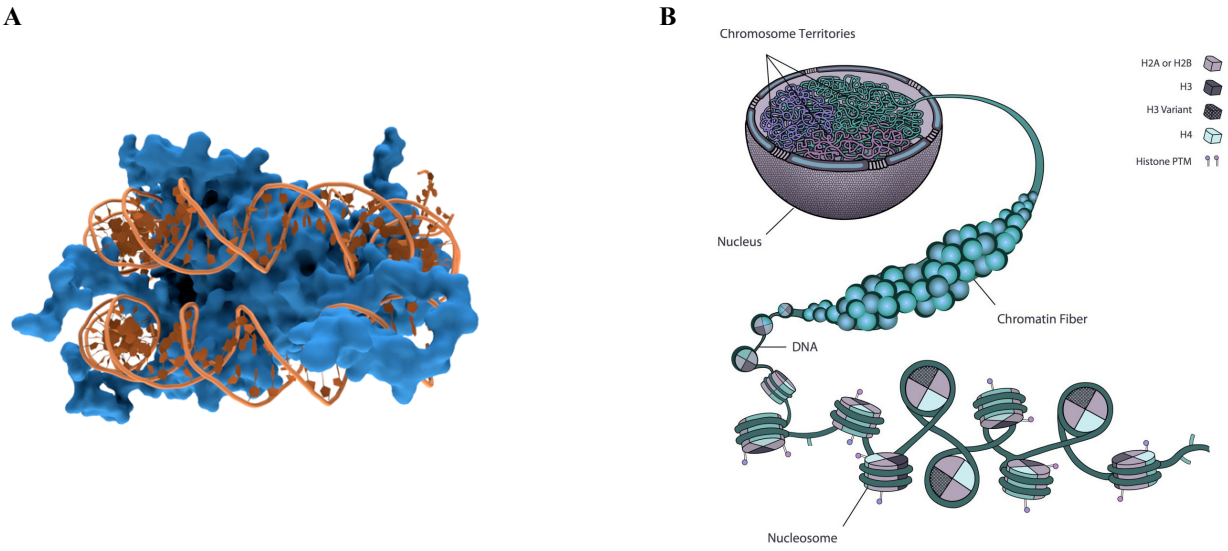


Figure 1.1 Hierarchical organization of the eukaryotic genome
A) DNA (orange) is wrapped around a histone octamer (blue) to generate nucleosome units. B) Nucleosomes are packaged into chromatin fiber, arranged into chromosomes and stored in the nucleus. Schematic illustrates the occurrence of post-translational modifications at specific histone tails. Adapted from (Rosa and Shaw, 2013). Illustration of nucleosome structure by Thomas Splettstoesser (CC-BY-SA-3.0).

Covalent modifications of these unstructured regions provide a fundamental mechanism that regulates chromatin dynamics and is responsible for faithful execution of critical biological programs (Kumar et al., 2016). Not surprisingly, a large number of histone modifying enzymes have been identified and are known to be responsible for a wide variety of histone post-translational modifications including acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation (Sadakierska-Chudy and Filip, 2014).

Opposite to the molecules that deposit histone modifications are the enzymes that remove or modify histone marks (i.e. histone deacetylases and lysine specific demethylases). Often, the "writers" and "erasers" of histone marks are found engaged in dynamic chromatin remodeling events at the same genomic locations. A third functional class of histone interacting proteins, also referred to as "readers", are in charge of decoding the epigenetic information established through the combination of histone marks. Once the decoder proteins recognize and bind specific histone modifications, they either modify the associated chromatin for subsequent DNA

interactions or directly recruit critical machinery for regulation of biological processes (Yun et al., 2011).

The balance between deposition and removal of histone modifications plays a critical role in the regulation of gene expression and maintenance of cellular homeostasis. For instance, altered expression of genes that encode histone deacetylase enzymes have been linked to the development of cancer and other diseases, as deacetylases are known to influence the transcription of key genes involved in cell proliferation, cell-cycle regulation, and apoptosis (Falkenberg and Johnstone, 2014). In the following section I will provide a description of specific protein factors, multi-subunit complexes, and known mechanisms that tightly regulate post-translational modification of histones. In addition, I will also discuss the currently known functions of specific covalent histone modifications primarily focusing on gene expression regulation.

1.2 Chromatin remodelers

In general, proteins that regulate the accessibility of chromatin are categorized into two main groups: enzymes that covalently modify histone proteins (histone modifying enzymes) and those that are dependent on ATP hydrolysis to change the positions of or evict histones from DNA (ATP-dependent chromatin-remodeling factors). ATP-dependent chromatin-remodeling factors are multi-protein complexes that harbor ATPase subunits. Four main classes of ATPase subunits have been characterized: SWI/SNF, ISWI, CHD and INO80 (Narlikar et al., 2013). Chromatin remodeling complexes have been implicated in a wide variety of cellular processes including facilitation and repression of gene expression, nuclear organization, centromere function, as well as chromosomal stability (Tsukiyama, 2002). In agreement with this broad functionality, a genome wide analysis found the SWI/SNF complex in close proximity to *cis*-

regulatory elements integral to transcription (i.e. enhancers), as well as occupying regions critical for chromosomal organization (i.e. DNA replication origins) (Euskirchen et al., 2011).

ATP-dependent chromatin-remodeling factors are most often responsible for the relaxation of chromatin. In contrast, histone modifying enzymes can induce both open chromatin (euchromatin) or repressed chromatin (heterochromatin) states. Covalent modifications of histones can change chromatin structure by targeting the interaction between DNA and histone residues. For instance, acetylation of lysines on histone tails neutralizes the positive charge of the targeted residue and reduces the affinity of histones for the negatively charged DNA backbone, thereby loosening the structure of the nucleosome (Smith and Denu, 2009). This enables the transcriptional machinery to access the DNA and enhances gene transcription (Kouzarides, 2007). Conversely, enzymatic removal of the acetyl groups from histone tails reverses the effect and leads to the formation of compact and transcriptionally repressed chromatin (Chi et al., 2010).

Together, ATP-dependent factors and histone modifying enzymes provide eukaryotic cells with a large repertoire of specialized chromatin remodeling units. Histone remodeling factors have evolved to target specific genomic features and help execute critical biological processes during different developmental stages (Clapier and Cairns, 2009). The following discussion is focused on histone modifying enzymes, since these proteins are more pertinent to my studies.

1.2.1 Histone methyltransferases (HMTs)

Methylation of histones is mediated by enzymes known as histone methyltransferases (HMTs). Most of our knowledge about histone methylation comes from studies of the *Saccharomyces cerevisiae* Set1 protein. Set1 is found within a multiprotein complex known as

COMPASS (complex of proteins associated with Set1); as a whole, these proteins are responsible for the mono-, di-, and trimethylation states of histone H3 on lysine 4 (H3K4) (Miller et al., 2001; Roguev et al., 2002). The enzymatic activity of Set1 is represented in *Drosophila melanogaster* by three related H3K4 methylases: dSet1, Trithorax (Trx), and Trithorax-related (Trr) (Mohan et al., 2011). In the mammalian system, there is a total of six COMPASS-like complexes: SET1A, SET1B, MLL1, MLL2, MLL3, and MLL4 (Shilatifard, 2012).

It appears that the reason for having multiple histone methylases originates both from the need to modulate the methylation levels on specific histone residues and the need to target distinct chromatin features for histone methylation (Table 1.1). All COMPASS family members from yeast to human possess a common protein core, including the methyl-transferase C-terminal SET domain or Set1/MLL, Cps60/Ash2L, Cps50/RbBP5, Cps30/WDR5, and Cps25/Dpy30 (Takahashi et al., 2011). However, unique subunit components are likely responsible for selective histone methylation and genomic feature deposition. According to this model, *Drosophila* Set1 and its mammalian homologs, SET1A and SET1B, have been recently shown to be responsible for the bulk level of di- and trimethylation of H3K4 (Hallson et al., 2012). Conversely, MLL1 is minimally required (~5%) for H3K4me3 at promoters of genes in mouse embryonic fibroblasts (MEFs) (Wang et al., 2009a). In *Drosophila*, Trr was shown to be responsible for monomethylation of H3K4 at *cis*-regulatory elements and the mammalian homolog MLL4 was also primarily enriched at enhancer regions in human colon cancer (HCT116) cells (Herz et al., 2012; Hu et al., 2013). Recent genome-wide analyses strongly suggest that in mammalian cells, particular chromatin context dictates the type of histone methyltransferase usage. For example, in mouse embryonic stem (ES) cells, MLL2 is the only

enzyme known to trimethylate H3K4 at bivalent genes (Deqing et al., 2013).

Table 1.1 Major histone acetyltransferases and methyltransferases
TA: transcriptional activation, TR: transcriptional repression, RA: DNA repair, RE: DNA replication. Adapted from (Kim, 2014).

Modification	Histone	Residue	Enzyme	Possible role
Acetylation	H2A	K4	Esa1	TA
		K5	Tip60, Hat1, p300/CBP	TA, RA
		K7	Hat1, Esa1	TA
	H2B	K5	ATF2	TA
		K11	Gcn5	TA
		K12	ATF2, p300/CBP	TA
		K15	ATF2, p300/CBP	TA
		K16	Gcn5, Esa1	TA
		K20	p300	TA
	H3	K4	Esa1, Hpa2	TA
		K9	Gcn5, SRC-1	TA, RA
		K14	Gcn5, PCAF, Tip60, SRC-1, hTFIIIC90, TAF1, p300/Gcn5, Esa1, Elp3, Hpa2, TAF1, Sas2, Sas3	TA, RA, RE
		K18	p300, CBP/Gcn5 (SAGA)	TA, RA
		K23	p300, CBP/Gcn5 (SAGA), Sas3	TA, RA
		K27	Gcn5	TA, RA
	H4	K5	Hat1, Tip60, ATF2, p300/Hat1, Esa1, Hpa2	TA, RA, RE
		K8	Gcn5, PCAF, Tip60, ATF2, p300/Esa1, Elp3	TA, RA, RE
		K12	Hat1, Tip60/Hat1, Esa1, Hpa2	TA, RA, RE
		K16	MOF, Gcn5, Tip60, ATF2/Gcn5, Esa1, Sas2	TA, RA
Methylation	H1	K26	EZH2	TR
	H3	R2	CARM1	TA
		K4	MLL4, SET1, MLL1, SET7/9, MYD3/Set1	TA
		R8	PMRT5	TR
		K9	SUV39h1, SUV39h2, ESET, G9A, EZH2, Eu-HMTase1/Clr4, S.p. Clr4	TA, TR
		R17	CARM1	TA
		R26	CARM1	TA
		K27	EZH2, G9A	TA, TR
		K36	HYPB, NSD1/Set2, S.c.	TA
		K79	DOT1L/S.c. Dot-1	TA, TR, RA
	H4	R3	PRMT1, PRMT5	TA
		K20	PR-SET7, SUV4-20/SET9	TA, TR, RA

Furthermore, during macrophage differentiation there is a major epigenetic remodeling of enhancer regions and knockdown of MLL1 and MLL3, but not MLL4, results in a significant decrease in monomethylation of H3K4 at *de novo* enhancers (Kaikkonen et al., 2013).

Although there are other classes of HTMs, the six COMPASS-like complexes are responsible for most of the methylation on histone residues in mammalian cells. The variety of HTMs provides a means to selectively target specific genome features for histone methylation

and the ability to modulate both the frequency and the degree of the modification. Below, I provide a description regarding the specific residues targeted on histones for methylation and the associated biological implications of specific histone methylations.

1.2.2 Histone lysine demethylases (KDMs)

Histone methylation was originally believed to be a stable modification, but we now know that specialized proteins catalyze the removal of methyl groups from histones (Shi et al., 2004). Two families of demethylases have been identified: the lysine-specific demethylase (LSD) family of enzymes is composed of LSD1 and LSD2 proteins, and the JMJC family of demethylases which contain the Jumonji C (JMJC) domain. Enzymes from both families have been reported to dynamically regulate demethylation of histones and non-histone proteins (Kooistra and Helin, 2012; Tsukada et al., 2006).

Similar to other chromatin interacting enzymes, members of the LSD family of demethylases are found within multiprotein complexes that mediate location-specific DNA interactions. The LSD1 enzyme displays a broad target specificity and can catalyze the removal of methyl groups from H3K4me1 and H3K4me2 at enhancers as well as the demethylation of H3K9me1 and H3K9me2 at promoter regions (Metzger et al., 2005; Wissmann et al., 2007). LSD1 can also remove methyl groups from non-histone proteins, such as K370me1 and K370me2 on the tumor protein p53 (Huang et al., 2007). Not surprisingly, decreased activity of LSD1 has been shown to suppress cell proliferation and induce significant transcriptome changes in colon cancer cells (Shi et al., 2004).

There are 30 members of the JMJC family of demethylases, more than half of which target histone residues. A well-studied member of the JMJC family is the F-box and Leu-rich repeat protein 11 (FBXL11) which is responsible for the demethylation of H3K36me1 and

H3K36me2 and is essential for embryonic development (Kawakami et al., 2015; Tsukada et al., 2006).

Histone methylases and demethylases fulfill a key role in the regulation of critical cellular processes, specifically at the level of transcription modulation. Many studies imply that the association of histone demethylases with transcription co-regulators at promoter regions results in the removal of repressive methyl histone marks and consequently aid in the initiation of transcription (Metzger et al., 2005; Yamane et al., 2006). However, deactivation of functional enhancers during cell fate establishment, attributed to members of the LSD demethylase family, has been shown to decommission enhancers and consequently repress gene expression (Whyte et al., 2012). Together, these features show that a proper balance from the activities of both HTMs and KDMs is crucial for the expression and reduction of gene transcripts.

1.2.3 Histone acetyltransferases (HATs)

Among all the histone post-translational modifications, acetylation and deacetylation are relatively the best understood. The levels of histone acetylation are controlled by the opposing actions of two classes of enzymes: histone acetyl transferases (HATs), which catalyze the deposition of acetyl groups from the coenzyme acetyl CoA to the ε-amino group of lysine side chains, and histone deacetylases (HDACs), which remove the modification (Figure 1.2) (Ververis et al., 2013). One of the first observations linking histone modifications to transcriptional regulation was made upon the characterization of the *Tetrahymena* acetyltransferase A, the homolog to the yeast transcription factor Gcn5p (Brownell et al., 1996).

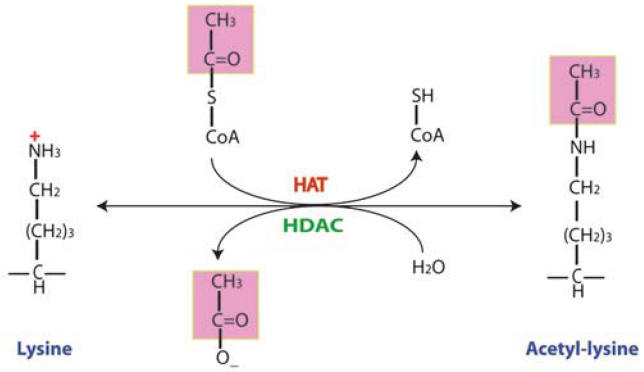


Figure 1.2 The process of enzyme mediated acetylation and deacetylation
Histone acetyltransferase enzyme (red) catalyzes the transferring of the acetyl moiety (pink) from acetyl-CoA to the ε-group of a lysine residue within the target protein. This even neutralizes the positive charge on the targeted lysine. The Zn^{2+} dependent classical HDAC (green) is responsible for the removal of the acetyl group from the acetylated-lysine. The reaction liberates acetate. Adapted from (Sun et al., 2012).

Subsequent studies later revealed that transcriptional events are greatly amplified though histone acetylation of promoter regions (Brown et al., 2000). Since then, many different HATs that function as transcriptional coactivators have been described.

HATs are categorized in two main families: type A and type B. Unlike the type A family of HATs, members of the type B family are predominantly cytoplasmic and acetylate newly synthesized histones that are not associated with chromatin. This type of acetylation is believed to play a major role in histone deposition and it is short lived, since the acetylation marks are removed soon after nucleosome formation (Parthun, 2007). Three main classes of type A HATs have been characterized: (1) the general control non-derepressible 5 (Gcn5)-related N-acetyltransferases known as the GNAT family, (2) the E1A-associated protein of 300 kDa (p300)/cAMP-responsive element binding proteins (CREB), and (3) the MYST proteins (Allis et al., 2007; Lee and Workman, 2007).

Members of the GNAT family share several sequence motifs that provide recognition of chromatin substrates and regulation of specific genes (Xu et al., 1998). In higher eukaryotes,

there are two Gcn5-like proteins, PCAF and GCN5. Both of these acetyl transferases can associate with p300/CBP proteins to form similar HAT complexes and are known to primarily influence transcription regulation and cell growth control (Yang et al., 1996). GCN5 mainly occupies promoter regions and has been implicated in displacement of promoter-associated nucleosomes prior to transcription activation (Barbaric et al., 2001). Moreover, GCN5 is believed to play an important role in the recruitment of multiple factors associated with gene expression such as RNA Polymerase II holoenzyme (RNA Pol II), the TATA box binding protein (TBP), and other general transcription coactivators (Qiu et al., 2004). GCN5 is mainly responsible for the acetylation of lysines 9, -14, -18, -23, and -27 on histone H3, and lysines 8 and -16 on histone H4 (Table 1.1).

The p300/CBP family consists of two transcription regulatory proteins that are highly homologous. Both p300 and CBP are broadly expressed across cell types and their functions are critical for the regulation of cell growth, differentiation, and apoptosis (Eckner et al., 1994). In vitro, recombinant p300/CBP molecules are capable of acetylating all four core histones both in their free form or when complexed with double stranded DNA. Indeed, p300 and CBP are highly efficient acetyl transferases and they both display broad substrate specificity. For instance, p300 is highly enriched at promoters and along *cis*-regulatory elements and can acetylate lysines 14, -18, -23, and -27 on histone H3 (Table 1.1).

Genome-wide studies have demonstrated the preferential enzyme-substrate interaction between p300 and H3K27. The dependency of H3K27ac on p300 catalytic activity was demonstrated in a recent study where knockdown of p300 in HCT116 cells resulted in the dramatic decrease of acetylated H3K27 (Tang et al., 2013). As discussed below, the presence of H3K27ac separates active from poised or disengaged enhancer regions (Bonn et al., 2012;

Creyghton et al., 2010; Zentner et al., 2011). This implies that among other functions p300 is the primary HAT responsible for maintaining the active state of the thousands of canonical enhancers.

1.2.4 HATs complexes

HATs do not function by themselves *in vivo*; instead, most HATs act in association with recruiting proteins or as components of large protein complexes. This type of protein cooperativity allows HATs to target a large number of diverse substrates and mediates histone acetylation at specific genomic locations.

There are four main HAT-associated multiprotein hubs that have been identified: the SAGA, PCAF, NuA3/4, and MSL complexes. Among them, the SAGA and PCAF complexes are the best understood. The SAGA complex, and the much smaller yeast homolog ADA, contain the GCN5 acetyl transferase and several isoforms of the adenosine deaminase proteins. In addition, the SAGA complex harbors at least five TATA box-binding associated factors (TAFIIs) and Sin4, a subunit of the RNA Pol II mediator complex. The SAGA complex is responsible for the recruitment of GCN5 to promoter regions for the acetylation of both H3K9 and H3K14 (John et al., 2000).

The human PCAF complexes are also similar to the yeast ADA complex in that they are both restrictively associated with the adenosine deaminase proteins Ada2 and Ada3 (Roth et al., 2001). For the most part, HAT-multiprotein complexes are composed of subunits displaying domains that mediate the recruitment of the HAT to specific genomic locations. Common protein domains found in these subunits include bromodomains (acetyl-binding), chromodomains (methyl-binding), WD40 repeats, Tudor domains and PHD fingers (Lee and Workman, 2007). Accordingly, a genome wide study looking at the global distribution of HATs

in $CD4^+$ T cells found that p300 and CBP were found enriched at enhancer-like and promoter regions whereas MOF, PCAF, and Tip60 were predominantly found at promoters and transcribable regions (Wang et al., 2009b). This suggests that particular incorporation of adapter subunits within HAT-complexes dictates the recruitment of distinct HATs at specific genomic locations to establish histone acetylation patterns.

1.2.5 Histone deacetylases (HDACs)

Opposing the enzymatic action of HATs are histone deacetylases (HDACs). HDACs catalyze the deacetylation of ε-acetyl-lysine residues from histones and other proteins (Figure 1.2). Consequently, HDACs are just as critical as HATs for proper cellular maintenance. Histone deacetylase enzymes influence the expression of genes involved in cell proliferation, cell-cycle regulation, and apoptosis, and their deregulation is often associated with tumor development (Falkenberg and Johnstone, 2014).

The HDAC superfamily has been categorized into four main classes based on their evolutionary conservation and functional properties. The classical zinc-dependent metalloproteins belong to classes I, II and IV. Class III houses the nonmetal-dependent enzymes known as the sirtuins (Yang and Seto, 2008). There are seven sirtuins (SIRT1-7), and their deacetylase activity is dependent on nicotinamide adenine dinucleotide (NAD^+) (Table 1.2). Members of the sirtuin family have been associated with cell proliferation and cell cycle control, but it appears that their primary function involves the regulation of metabolic homeostasis (Hall et al., 2013; Houtkooper et al., 2012). For example, SIRT3 is known as a major mitochondrial deacetylase and SIRT1 can promote the induction of mitochondrial gene expression as well as lipid and glucose metabolism (Bao et al., 2010; Lagouge et al., 2006).

Table 1.2 Classification histone deacetylase enzymes
Shown are the 18 histone deacetylase enzymes classifications, peptide length, cellular localization, and associated physiological function. Adapted from (Mottamal et al., 2015).

				Metal Dependent
Class	**Members**	**Size (aa)**	**Cellular Localization**	**Physiological Function**
I	HDAC1	483	Nucleus	Cell survival and proliferation
	HDAC2	488	Nucleus	Cell proliferation, Insulin resistance
	HDAC3	428	Nucleus	Cell survival and proliferation
	HDAC8	377	Nucleus	Cell proliferation
IIA	HDAC4	1084	Nucleus/Cytoplasm	Regulation of skeletogenesis and gluconeogenesis
	HDAC5	1122	Nucleus/Cytoplasm	Cardiovascular growth and function, gluconeogenesis, cardiac myocytes and endothelial cell function
	HDAC7	912	Nucleus/Cytoplasm	Thymocyte differentiation, endothelial function, glucogenesis
	HDAC9	1069	Nucleus/Cytoplasm	Homologous recombination, thymocyte differentiation, cardiovascular growth and function
IIB	HDAC6	1215	Cytoplasm	Cell motility, control of cytoskeletal dynamics
	HDAC10	669	Cytoplasm	Homologous recombination, Autophagy mediated cell- survival
IV	HDA11	347	Nucleus	Immunomodulators-DNA replication
				NAD+ Dependent
III	SIRT 1	747	Nucleus, Cytoplasm	Aging, redox regulation, cell survival, autoimmune system regulation
	SIRT 2	389	Nucleus	Cell survival-cell migration and invasion
	SIRT 3	399	Mitochondria	Urea Cycle, Redox balance, ATP regulation, metabolism, apoptosis and cell signaling
	SIRT 4	314	Mitochondria	Energy metabolism, ATP regulation, metabolism, apoptosis and cell signaling
	SIRT 5	310	Mitochondria	Urea cycle, Energy metabolism, ATP regulation, metabolism, apoptosis and cell signaling
	SIRT 6	355	Nucleus	Metabolic regulation
	SIRT 7	400	Nucleus	Apoptosis

However, results from a more recent study point to a direct link between sirtuins and the modulation of histone acetylation levels. Specifically, in normal rat cardiomyocytes, SIRT6 directly interacts with a c-Jun-containing AP-1 complex, is subsequently recruited to proximal promoter regions displaying the AP-1 DNA binding site and efficiently represses gene expression by deacetylating histone H3 at lysine 9 (Sundaresan et al., 2012).

The class I family of classical HDACs consists of HDAC1, 2, 3, and 8 enzymes, all sharing sequence homology with the yeast Rpd3 protein. Class I HDACs are primarily localized to the nucleus (Brosch et al., 2008; Gregoretti et al., 2004; de Ruijter et al., 2003). On the other hand, class II family HDACs are homologous to the yeast histone deacetylase 1 (Hda1) and are

further divided into two subgroups, class IIa (HDAC4, 5, 7, and 9) and class IIb (HDAC6 and 10) (Table 1.2). Unlike class I HDACs, class II HDACs are mainly localized to the cytoplasm. Nonetheless, the phosphorylation status for some class II HDACs can induce shuttling between the cytoplasm and the nucleus. The class IIb HDAC6 was recently found to associate with protein coding gene regions where it is likely responsible for the modulation of acetylated histone H3 (Wang et al., 2009b; Yang and Gregoire, 2005). HDAC11 is the only member of the class IV family and has only been detected in the nucleus. HDAC11 has a unique structure, but it also displays some sequence similarities to class I and II enzymes (Barneda-Zahonero and Parra, 2012).

Functional redundancy among different enzymes within class I HDACs makes the identification of specific targeted lysines on histones a difficult task. However, a recent investigation in which a double HDAC1/2 knockdown cell system was established for the first time, revealed that the two closely related metalloenzymes primarily target residues on histone H3 for deacetylation (Jamaladdin et al., 2014). In this system, acetylation of H3K14 and H3K56 was increased by over threefold while acetylation of lysines 9, 18, 23, and 27 on the same histone was modestly elevated. Of note, the levels of dimethylated H3K4 and trimethylated H3K9 did not show significant changes by immunoblotting, suggesting that class I HDACs are not responsible for large-scale remodeling of nucleosome methylation.

Besides the deacetylation of histones, classical HDACs also target non-histone proteins for the removal of acetyl groups. A surprising observation is that phylogenetic analyses show non-histones as the likely primary substrates of HDACs (Gregoretti et al., 2004).

Table 1.3 Non-histone protein substrates of classical HDACs
Abbreviations: SRY, sex-determining region Y; STAT, signal transducer and activator of transcription; GATA, GATA-binding factor; E2F, E2F transcription factor; MyoD, myogenic differentiation; YY1, transcriptional repressor protein; HMG, High Mobility Group; AR, androgen receptor; ER, estrogen receptor; EKLF, Erythroid Kruppel- like factor; RUNX, Runt-related transcription factor; HIF, Hypoxia-inducible factor; NF-kB, nuclear factor kappa-B; Ku70, ATP-dependent DNA helicase; Hsp, heat-shock protein. Adapted from (Ververis et al., 2013).

Effect of acetylation	Protein	Intracellular function	HDAC implicated
Increased DNA-binding affinity	p53	Tumor suppressor	
	SRY	Transcription factor	HDAC3
	STAT3	Signaling mediator	HDAC1, -2, -3
	GATA1	Transcription factor	HDAC3, -4, -5
	GATA2	Transcription factor	HDAC3, -5
	E2F1	Transcription factor	HDAC1
	MyoD	Transcription factor	HDAC1
Decreased DNA-binding affinity	YY1	Transcription factor	HDAC1, -2, -3
	HMG-A1	Nuclear factor	
	HMG-N2	Nuclear factor	
	p65	Transcription factor	
Increased transcriptional activation	p53	Tumor suppressor	
	HMG-A1	Nuclear factor	
	STAT3	Signaling mediator	HDAC1, -2, -3
	AR	Nuclear receptor	HDAC1
	ERα (basal)	Steroid hormone receptors	HDAC1
	GATA1	Transcription factor	HDAC3, -4, -5
	GATA2	Transcription factor	HDAC3, -5
	GATA3	Transcription factor	
	EKLF	Transcription factor	HDAC1
	MyoD	Transcription factor	HDAC1
	E2F1	Transcription factor	HDAC1
	RUNX3	Tumor suppressor	HDAC1, -5
Decreased transcriptional activation	ERα (ligand-dependent)	Steroid hormone receptors	HDAC1
	HIF1α	Transcription factor	
Increased protein stability	p53	Tumor suppressor	HDAC1
	c-MYC	Oncoprotein	
	AR	Nuclear receptor	HDAC1
	ERα	Steroid hormone receptors	HDAC1
	E2F1	Transcription factor	HDAC1
	Smad7	Signaling mediator	HDAC1, -3
	RUNX3	Tumor suppressor	HDAC1, -5
Decreased protein stability	HIF1α	Transcription factor	
Promotes protein–protein interaction	STAT3	Signaling mediator	HDAC1, -2, -3
	AR	Nuclear receptor	HDAC1
	EKLF	Transcription factor	HDAC1
	Importin α	Nuclear import factors	
Disrupts protein–protein interaction	NF-κB	Transcription factor	
	Ku70	DNA-repair protein	
	Hsp90	Chaperone	HDAC6

Indeed, HATs and HDACs can influence gene expression by directly modulating the functionality of transcription factors and other regulatory proteins (Ocker, 2010). There are more than fifty non-histone proteins subject to specific deacetylation by classical HDACs, including transcription factors (i.e. p53, NF-κB, and E2F1), transcriptional co-regulators (Rb),

enzymes involved in DNA repair (Ku70 and NEIL2), structural proteins (α-tubulin) and chaperones (Hsp90) (Table 1.3) (Ververis et al., 2013; Xu et al., 2007). Consequently, there is a broad cellular impact resulting from HDAC modulation of non-histone proteins, ranging from mRNA instability to cytoskeleton dynamics. For instance, HDAC6 has been shown to regulate α-tubulin deacetylation and it is found to associate with microtubules. Ectopic expression of HDAC6 usually promotes chemotaxis, suggesting that HDAC6 regulates microtubule-dependent cell motility (Hubbert et al., 2002).

1.2.6 Class I HDAC complexes

On their own, HDACs display low substrate selectivity and *in vitro* analysis shows that a single enzyme can deacetylate multiple sites within a target protein (Riester et al., 2007). In higher eukaryotes the substrate selectivity is likely more restricted due to the incorporation of HDAC proteins into specialized multiprotein complexes. With the exception of HDAC8, all class I members can function as catalytic subunits. In mammals, HDAC1 and HDAC2 coexist as the catalytic core of three multi-subunit complexes. The double HDAC catalytic core is present in the mammalian Sin3, in the nucleosome remodeling deacetylases (NuRD), and in the corepressor of RE1-silencing transcription factor (CoREST) complexes (Table 1.4) (Grozinger and Schreiber, 2002).

All three of these complexes have been shown to cooperate with other chromatin remodelers in a DNA-sequence specific manner to repress transcription. In addition to the HDACs, these complexes carry other catalytic subunits with histone remodeling properties that often complement deacetylase function. For instance, the CoREST-HDAC associated complex harbors the histone H3K4 demethylase LSD1.

Table 1.4 Class I HDAC complexes and associated subunits
Shown are the subunits of mammalian HDAC complexes characterized through biochemical purifications. Adapted from (Yang and Seto, 2008).

Complex	Mammals	Protein domain
Sin3	HDAC1, HDAC2	Class I deacetylase
	RbAp46, RbAp48	WD40 repeat
	Sin3A	PAH motifs
	Sds3, BRMS1	Coiled coil
	RBP1	
	SAP30	
	SAP18	Ubiquitin fold
	ING1/2	PHD finger
Mi-2/NuRD	HDAC1, HDAC2	Class I deacetylase
	RbAp46, RbAp48	WD40 repeat
	Mi-2α/β	Helicase
	MTA1–3	SANT domain
	MBD2, MBD3	Methyl CpG binding
	p66α/β	
CoREST	HDAC1, HDAC2	Class I deacetylase
	CoREST	SANT domain
	LSD1	SWIRM, K-demethylase
	BHC80	PHD finger
	CtBP1	Dehydrogenase
N-CoR/SMRT	HDAC3	Class I deacetylase
	N-CoR/SMRT	SANT domain
	TBL1/TBLR1	WD40 repeat
	GPS2	
	JMJD2A	Jumonji demethylase
	Kaiso	Methyl CpG binding

It has been suggested that HDACs and LSD1 are functionally co-dependent or may work synergistically in the removal of acetyl and methyl groups to generate a repressive chromatin environment (Shi et al., 2005). A recent study found that LSD1 and HDAC1 and HDAC2 act cooperatively at active enhancers to remove H3K4me1 and H3K27ac upon differentiation of mouse embryonic stem (ESCs) cells, allowing for the inactivation of pluripotency genes (Whyte et al., 2012).

Ultimately, the dynamic cycle of histone post-translational modifications and the removal of such marks (i.e. acetyl or methyl groups) by the remodeling enzymes provide indispensable management of chromatin accessibility. Because histone remodeling enzymes are often part of multi-subunit complexes, the accessory proteins in these hubs dictate histone remodeling patterns and influence the targeting of specific genomic locations to activate or repress gene transcription.

1.3 Covalent histone modifications and associated functions

Histone tails are targeted by a variety of posttranslational modifications, among them acetylation and methylation. More often than not, multiple modifications exist on any given histone protein. The combination of these modification patterns prompted the hypothesis about of a "histone code". The histone modifications are highly dynamic as a result of competing actions of histone mark "erasers" such as KDMs or HDACs with histone mark "writers" like histone methyltransferases or histone acetylases. The interplay between writers, erasers and readers gives rise to distinct chromatin configurations which are ultimately responsible for triggering a diverse array of biological responses in a context-dependent manner.

1.3.1 Histone acetylation

Histone acetylation is best understood for its localization at proximal promoters and positive contribution to gene expression. However, plenty of evidence gathered through the past decade has established that specific histone acetylation marks are indispensable for maintaining enhancer elements in a functional state (Bonn et al., 2012; Creyghton et al., 2010; Zentner et al., 2011). In general, acetylation of lysines on histone tails which occurs at higher frequencies on histone H3, positively contributes to transcription activation but the deposition of the functional group has also been associated to histone deposition, DNA repair, and telomeric silencing (Table 1.5) (Rea et al., 2000; Zhang et al., 2002).

Table 1.5 Covalent modifications of histone tails and associated cellular function
Adapted from (Lawrence et al., 2016).

Histone	Modification	Role
H2A	H2AS1P	Mitosis; chromatin assembly
	H2AK4/5ac	Transcriptional activation
	H2AK7ac	Transcriptional activation
	H2AK119P	Spermatogenesis
	H2AK119uq	Transcriptional repression
H2B	H2BS14P	Apoptosis
	H2BS33P	Transcriptional activation
	H2BK5ac	Transcriptional activation
	H2BK11/12ac	Transcriptional activation
	H2BK15/16ac	Transcriptional activation
	H2BK20ac	Transcriptional activation
	H2BK120uq	Spermatogenesis/meiosis
	H2BK123uq	Transcriptional activation
H3	H3K4me2	Permissive euchromatin
	H3K4me3	Transcriptional elongation; active euchromatin
	H3K9me3	Transcriptional repression; imprinting; DNA methylation
	H3R17me	Transcriptional activation
	H3K27me3	Transcriptional silencing; X-inactivation; bivalent genes/gene poising
	H3K36me3	Transcriptional elongation
	H3K4ac	Transcriptional activation
	H3K9ac	Histone deposition; transcriptional activation
	H3K14ac	Transcriptional activation; DNA repair
	H1K18ac	Transcriptional activation; DNA repair; DNA replication
	H3K23ac	Transcriptional activation; DNA repair
	H3K27ac	Transcriptional activation
	H3T3P	Mitosis
	H3S10P	Mitosis; meiosis; transcriptional activation
	H3T11/S28P	Mitosis
H4	H4R3me	Transcriptional activation
	H4K20me1	Transcriptional silencing
	H4K20me3	Heterochromatin
	H4K5ac	Histone deposition; transcriptional activation; DNA repair
	H4K8ac	Transcriptional activation; DNA repair; transcriptional elongation
	H4K12ac	Histone deposition; telomeric silencing; transcriptional activation; DNA repair
	H4K16ac	Transcriptional activation; DNA repair
	H4S1P	Mitosis

Acetylation of histone H3 on lysine 27 (H3K27ac) has been widely associated with active promoters in mammalian cells (Wang et al., 2008) and can be deposited by both p300 and CREB binding protein (Tie et al., 2009). High levels of H3K27ac modification at promoter regions have been reported to aid in the recruitment of RNA Pol II and promote the transition from initiation into the elongation state (Stasevich et al., 2014). Importantly, H3K27ac has emerged as the hallmark of active enhancers and is predominantly seen partnered with H3K4me1 at *cis*-regulatory regions (Bonn et al., 2012; Creyghton et al., 2010; Rada-Iglesias et al., 2011; Zentner et al., 2011). In contrast to H3K27ac, H3K4me1 can also occur at inactive but primed enhancers (Hu et al., 2013; Kaikkonen et al., 2013). As mentioned above, the interdependence of these two histone marks has been described in several model systems. For instance, during macrophage differentiation H3K27ac is deposited at genomic regions that have been previously marked with H3K4me1 to generate new enhancers (Bonn et al., 2012; Kaikkonen et al., 2013) . The tight association between these two histone marks is further underlined by evidence that shows a significant decrease in H3K27ac levels upon guided recruitment of the LSD1 histone demethylase to an endogenous enhancer (Forneris et al., 2005; Mendenhall et al., 2013).

Similar to H3K27ac, H3K9ac has been consistently associated with euchromatin and active gene expression. Not surprisingly, decline in H3K9ac accumulation has been observed upon chromatin stress in human cells, possibly to repress expression of genes associated with growth and cell cycle progression (Tjeertes et al., 2009; Zhou et al., 2010). In mouse embryonic stem cells, H3K9ac and H3K14ac show a high overlapping distribution, mainly along active promoters with high GC content. However, the same two histone marks can be detected associated with the repressive mark, H3K27me3, along transcriptionally inactive regions known as bivalent genes (Karmodiya et al., 2012). Less is known about the association of H3K9ac with

enhancer elements, but the mark has been used in recent genome-wide studies aiming to predict the location of active enhancers (Zhu et al., 2013). Acetylated H3K4 is mostly enriched at promoters of actively transcribed genes and is positioned just upstream of H3K4me3 (Guillemette et al., 2011).

1.3.2 Histone methylation

As mentioned above, acetylation of lysine residues in histone tails is more commonly associated with open chromatin and broadly accepted as an indicator of active regulatory elements and positive gene expression. On the other hand, complex patterns of histone lysine methylation encode distinct functions and can induce open or repressive chromatin (Table 1.5). Methylation marks tend to be versatile and dynamic with respect to gene activity. For example, di and tri-methylation of lysine 9 on histone H3 (H3K9me2/3) can occur not only at silent heterochromatin but also at the transcribed region of active mammalian genes (Vakoc et al., 2005). In the same manner, histone H3K4 can be mono-, di-, or tri-methylated. Mono-methylation of histone H3 on lysine 4 (H3K4me1) is found in a characteristic bifurcation shape flanking the transcription start site of actively transcribed or poised genes. More recently, a large number of genome-wide studies have consistently shown H3K4me1 to be also enriched at enhancers (Bulger and Groudine, 2011; Heintzman et al., 2007; Zhu et al., 2013). Accumulating evidence suggests that H3K4me1 is critical for the assembly of new enhancers that are deployed during the process of development and cell differentiation (Kaikkonen et al., 2013; Seumois et al., 2014).

Similar to H3K4me1, di-methylation of histone H3 on lysine 4 (H3K4me2) is also found at promoters and enhancers (Pekowska et al., 2011). However, this histone mark is not essential for the generation of active enhancers. Instead, it appears that H3K4me2 provides enhancer

stability or may even act to establish genomic memory. For example, during differentiation of macrophages where new enhancers are established, H3K4me2 is positioned hours after H3K27ac levels become significant and the synthesis of eRNA is apparent (Kaikkonen et al., 2013).

Histone H3K4 trimethylation (H3K4me3) is a prominent feature of promoters from actively transcribed genes but can also be found at poised or lowly transcribed genes co-occupied with H3K27me3, known as bivalent genes (Barski et al., 2007; Bernstein et al., 2006; Guenther et al., 2007). New evidence suggests that the presence of H3K4me3 at enhancers reflects the activity strength of the regulatory element as well as the high transcript levels of the corresponding gene (Core et al., 2014; Pekowska et al., 2011). While all three methylation states of H3K4 appear to positively contribute to active enhancers, the methylation of other lysine residues is likely to be context dependent. For example, methylation at H3K36 has a positive effect when it is found on the coding region of a gene but a negative effect when found at promoters (Vakoc et al., 2005).

1.4 Transcription of protein coding genes

Transcription is the process of RNA synthesis and results in the transfer of the information coded within the template DNA. The cycle of eukaryotic transcription can be generally partition in three main stages: initiation, elongation, and the processing of nascent transcript maturation. Each one of these steps is coordinated by a large number of factors and consequently transcription can be tightly regulated. Transcription factors at enhancers and promoters heavily modulate the levels of expressed RNA through the recruitment of both activating and repressive cofactors (Bulger and Groudine, 2011; Shlyueva et al., 2014; Spitz and Furlong, 2012). In the next section, I present a summary of the most relevant stages and protein factors involved in the process of transcription initiation, elongation and termination and a brief

description of enhancer elements and their influence on gene expression.

1.4.1 Transcription initiation

The initial synthesis of RNA transcripts is dependent on the recruitment and assembly of the pre-initiation complex (PIC) at the proximal promoter, including RNA Pol II, a battery of general transcription factors (GTFs) and the Mediator complex (Core et al., 2008; Lee and Young, 2000; Thomas and Chiang, 2006). The Mediator multi-subunit complex integrates activation and inhibition signals from transcription factors and coactivators around the RNA Pol II holoenzyme, ultimately regulating RNA Pol II-dependent transcription of each gene (Malik and Roeder, 2010; Taatjes, 2010). A key feature of the Mediator complex lies on the many subunits it harbors, which allow Mediator to specifically interact with a wide variety of sequence specific transcription factors, the subunits of the PIC (TBP, TFIIA, B, D, E, F and H and RNA Pol II). After establishment of the PIC at the transcription start site (TSS), short transcription pulses are generated ranging from ~20-60 bases before stalling. The pausing of RNA Pol II just downstream of the TSS provides a major regulatory feature in gene expression, allowing for priming of genes and a rapid deployment of transcripts in response to biological cues (Adelman and Lis, 2012; Gilchrist et al., 2010; Rahl et al., 2010). The synthesis of full RNA transcripts requires the release of paused RNA polymerase which is mediated by the presence of both positive regulatory proteins and specific histone modifications. More recently, distal enhancers have been implicated in the recruitment of such pause release factors.

1.4.2 RNA Pol II pausing and transcription elongation

Pausing of RNA Pol II at proximal promoters has been suggested to be a global event that regulates the majority of protein coding genes (Adelman and Lis, 2012). The positive transcription elongation factor (P-TEFb) is responsible for the release of RNA Pol II into the

gene body, which results in the production of full length pre-mRNA. P-TEFb is a heterodimer, consisting of the cyclin-dependent kinase CDK9 and its partner Cyclin T1 (CCNT1). The complex exists in two forms, as either an active dimer or an inactive structure when bound by both the inhibitory protein HEXIM1 and the single-stranded noncoding RNA 7SK (Byers et al., 2005). The active P-TEFb complex promotes transcriptional elongation by phosphorylating the paused-inducing factors DSIF and NELF, as well as serine 2 on the C-terminal domain of RNA Pol II (Kwak and Lis, 2013; Peterlin and Price, 2006). P-TEFb recruitment and catalytic activity is heavily dependent on the bromodomain-contining protein BRD4 (Moon et al., 2005; Yang et al., 2005). It has been suggested that all active P-TEFb kinase is found associated with BRD4, which highlights the critical role that BRD4 plays on transcription elongation (Moon et al., 2005).

BRD4 closely associates with the Mediator complex via several interacting domains and negotiates RNA Pol II pause release by liberating P-TEFb from the inhibitory factors, HEXIM1 and 7SK snRNA (Moon et al., 2005; Yang et al., 2005). Direct interaction between BRD4 and Cyclin T1 results in the transition of the inactive P-TEFb complex into the active state (Yang et al., 2005; Zhang et al., 2012). The two bromodomains in BDR4 selectively bind to acetylated residues on multiple proteins, including histones (Huang et al., 2009a; LeRoy et al., 2008). Not surprisingly, BRD4 has been reported to accumulate at enhancer elements and promote enhancer RNA transcription, particularly at cluster of enhancers co-occupied by the Mediator complex (Kanno et al., 2014; Lovén et al., 2013; Nagarajan et al., 2014; Zhang et al., 2012). Accordingly, a recent study has shown that BRD4 and the arginine demethylase JMJD6 can form a catalytic complex to activate P-TEFb and regulate the RNA Pol II pause release of a large number of genes. However, the BRD4/JMJD6 complex is not recruited directly to the proximal promoter

regions of genes with paused RNA Pol II, instead the functional association of BRD4 and JMJD6 is established at distal enhancer elements, referred to as "anti-pause enhancers" (Liu et al., 2013a).

Specifically, the BRD4/JMJD6 complex accumulates at distal enhancers enriched with H3K27ac, H3K4me1/2, and p300, all features of active enhancers. DNA looping facilitates the juxtaposition of the BRD4/JMJD4 complex to RNA Pol II paused promoters where JMJD6 targets both the repressive histone mark H4R3me2 and the methyl-cap of the 7SK shRNA for demethylation. This event results in the release of the pause inducing factors HEXIM1 and 7SK shRNA and activation of the P-TEFb complex (Figure 1.3). The regulation of RNA Pol II pause release is dynamic and critical for the maintenance of proper levels of gene expression. As discuss bellow, emerging evidence suggest that distal enhancer elements play a major role in the regulation of transcription initiation and elongation.

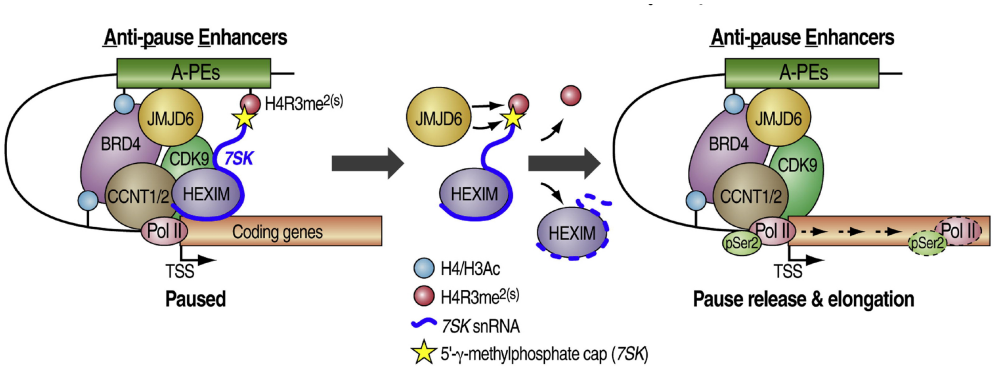


Figure 1.3 Regulation of RNA Pol II pause release coordinated by distal enhancers
Schematic diagram shown illustrates the mechanistic model for RNA Pol II pause release. Elongation inhibitors HEXIM and 7SK snRNA repress the kinase activity of the P-TEFb complex (CCNT1/2 and CDK9). Acetylated histones H3 and/or H4 at distal enhancer elements recruit the bromodomain-containing protein BRD4 associated to JMJD6. DNA looping facilitates the juxtaposition of the BRD4/JMJD6 protein complex to RNA Pol II paused promoters where both demethylases target H4R3me2 and 7SK snRNA methyl group in its cap structure. Release of the inhibitory factors HEXIM and 7SK snRNA permits the activation of P-TEFb and subsequent phosphorylation of RNA Pol II Ser2 for transcription elongation. Adapted from (Liu et al., 2013a).

Once the paused RNA Pol II is released from the proximal promoter, the holoenzyme enters the elongation stage and the full gene is transcribed. Nascent RNA is processed into a mature transcript as it is being transcribed during the elongation stage. This process entails specific modifications at both 5' and 3' ends and the removal of introns. Capping occurs at the 5' ends of RNAs through the addition of methylated guanosine triphosphate; this modification is necessary for subsequent processing and nuclear export (Lee and Young, 2000). Concomitantly, trimming and polyadenylation at the 3' end of RNAs takes place. The polyadenylation signal stimulates both the process of transcript termination and the release from the template genomic DNA and also provides stability to the transcript. Trimming of the transcript at the polyadenylation site is mediated by the cleavage stimulation factors, CPSF and CstF (Kuehner et al., 2011). Several studies have described the direct role of the RNA Pol II CTD as well as other components of the transcription apparatus in the process of nascent RNA capping, termination, and splicing (Hsin and Manley, 2012). Indeed, transcription by RNA Pol II assemblies with defective CTDs result in the synthesis of RNA molecules that are not capped (Lee and Young, 2000). Taken together, it appears that the mechanisms and molecular players involved in the regulation of the different stages of transcription are specific for distinct classes of genes, *cis*-regulatory elements such as enhancers provide an additional layer of transcriptional control.

1.5 Transcriptional enhancers

Although the core promoter and basal transcription factors are sufficient to initiate the process of transcription at a given gene, enhancers can greatly influence the level of gene expression (Juven-Gershon and Kadonaga, 2010). More importantly, enhancer elements serve as decoders of transcription factor inputs and together control the spatio-temporal distribution of gene expression and consequently cell fate decisions. Our understanding of enhancer elements

has grown exponentially over the last decade. With the advent of Nextgen sequencing technology, we are now capable of routinely analyzing networks of stimulated normal and transformed cells at a genome-wide level. In the following section, I will provide a brief overview of the findings on transcriptional enhancers most pertinent to this work.

1.5.1 *Cis*-regulatory elements (enhancers)

Early studies discovered that enhancer regions contain high GC-rich content, are sensitive to DNase I cleavage, and suggested their association to transcription regulation (Weintraub and Groudine, 1976). New genome wide technologies continue to reveal deeper information about enhancers, ranging from the identity of numerous associated histone marks to the molecular mechanisms that promote their establishment and decommission across many cells and tissue types. Estimates indicate that there may be ~40 thousands active enhancers in a functional mammalian cell and over a million enhancer-like regions encoded in the human genome (Andersson et al., 2014; Bernstein et al., 2012; Thurman et al., 2012). Better understanding of enhancer functional states and regulation is critical to understanding human development as well as disease.

The first enhancer identified was isolated from the SV40 viral genome and was capable of inducing a two hundred-fold increase in the expression of a rabbit beta-globin gene in HeLa cells (Banerji et al., 1981). The plasticity of the viral regulatory element became evident by its ability to retain influence on transcription events regardless of the distance or orientation in relationship to the reporter gene (Banerji et al., 1981). Soon after, enhancer elements were identified in more complex organisms including mammals (Banerji et al., 1983). A mouse enhancer located within the immunoglobulin heavy chain locus (IGH) provided the first evidence for cell type specificity: the enhancer functioned in myeloma cells (a lymphocyte-derived

tumor), but not in HeLa cells (derived from a cervical carcinoma) (Banerji et al., 1983; Davidson et al., 1986). Subsequently, it became evident from several reports that the activity of enhancers could not only be cell type-specific but also restricted to different developmental stages (Choi and Engel, 1986; Hesse et al., 1986; Kollias et al., 1987; Trudel and Costantini, 1987). Further studies indicated that the function of enhancers was dependent on the binding of *trans*-acting protein factors. Primarily, DNase I footprinting protection assays showed that each cell type contained different protein factors that bound to specific sequences encoded within the enhancer DNA (Augereau and Chambon, 1986; Mercola et al., 1985; Scholer and Gruss, 1985; Sen and Baltimore, 1986).

1.5.2 Transcription factors

A defining feature of enhancer elements is that they bind sequence-specific transcription factors. Enhancer regions, which range from 0.5 to 10 kilobases, carry multiple recognition motifs for the specific interaction with distinct classes of transcription factors. DNA binding sites are often evolutionarily conserved and permit the execution of cell type-specific gene expression programs (reviewed in (Levine, 2010)). In addition to the canonical motifs, enhancers also carry secondary recognition elements that mediate the interaction with additional regulatory proteins and molecular complexes, such as those for HATs and HDACs (Wang et al., 2012).

The recruitment of multiple proteins at a single element is another key feature of enhancers, highlighting the scaffold nature of enhancers. This process promotes cooperative binding and increases protein affinity to enhancer regions. In many cases, the first recruited protein can directly or indirectly facilitate the binding of additional co-regulators by either serving as a binding platform or by altering nucleosome conformation for additional recruitment.

The synergistic interactions of multiple *trans*-regulatory factors often results in a greater transcriptional output that could not be achieved by the action of a single factor (Carey et al., 1990; Giese et al., 1995; Kim and Maniatis, 1997).

The current estimate is that human cells contain the blueprints for ~1,500 transcription factors (Wingender et al., 2015). Yet, recent evidence shows that during differentiation, a small number of lineage-determining transcription factors act collectively to activate a large number of enhancers and mediate the expression of thousands of target genes (Heinz et al., 2010). This hierarchical regulation of transcription factors ultimately allows for the detrimental consequences observed in pathophysiological states, where perturbation of the function of a few key transcription factors can elicit a large variety of cellular and tissue responses (Hanahan and Weinberg, 2011).

1.5.3 DNA looping

Genome-wide location analysis of most signal dependent transcription factors indicate that the vast majority of their binding sites are in distal intra- and intergenic locations (Barish et al., 2010; Carroll et al., 2006; Chavanas et al., 2008). Unlike promoters, enhancers can regulate genes from long distances, generally under 100 kilobases, but in some cases from up to a megabase apart; such is the case of the highly conserved enhancer that regulates the Sonic hedgehog gene (Lettice et al., 2003; Sagai et al., 2004). Although enhancers can act upon nearby genes, it is estimated that only five percent of enhancers contribute to the transcription of their adjecent genes. Communication between distal enhancer regions and the core promoter is achieved through a "DNA looping mechanism". This event was first described in prokaryotic systems (Matthews, 1992; Ptashne, 1986). For example, the nitrogen regulatory protein C (NtrC) binds to an upstream enhancer far from the glnA operon and positively influences

transcription by forming a DNA loop (Su et al., 1990). This coupling mechanism is broadly used in eukaryotic systems and several studies have demonstrated the physical looping interactions between many individual enhancers and promoters in mammalian cells, including the β-globin enhancers and promoters (Tolhuis et al., 2002). The juxtaposition of distal enhancers with promoters is critical for gene activation. For example, targeted artificial looping was sufficient to activate gene expression at the β-globin locus (Deng et al., 2012). The functional importance of chromatin looping interactions was further highlighted with the discovery of enhancer-promoter insulators. Insulators are specialized DNA sequences that hinder enhancer-activated transcription only when found between an enhancer and its target promoter (Eissenberg and Elgin, 1991). For example, the *Drosophila* bithorax complex contains several insulators that confine both activator and repressor signals to specific regulatory domains, thereby maintaining proper gene expression along the anterior-posterior axis of the developing embryo (Galloni et al., 1993; Gyurkovics et al., 1990; Hagstrom et al., 1996).

Mediator and cohesin complexes are largely responsible for the structural stabilization of enhancer-promoter coupling. The cohesin complex, which is also involved in sister chromatid attachment, is a ring-shaped structure. While the precise details are still unknown, cohesin possibly facilitates looping by encircling the nucleosome-occupied enhancer and securing it near the Mediator complex (Guillou et al., 2010; Kagey et al., 2010). The loading of cohesin onto DNA is mediated by NIPBL, a protein that is tightly associated with the Mediator complex. Both cohesin and Mediator are known to co-occupy a large number of enhancers and promoters, likely participating in the selective association of genomic regions that promote gene specific transcription (Kagey et al., 2010).

Ultimately, the fidelity of transcription regulation relies on the specific interaction between transcription factors and enhancers. However, the promiscuous nature of enhancer activity, coupled with an ability to affect promoters from remote locations, makes the identification of the target genes a difficult task. Currently, global mapping of enhancer-promoter interactions is perhaps the biggest limiting factor in the field and greater ability to map chromatin interactions will improve our ability to understand how enhancers influence gene expression during development and disease.

1.5.4 Enhancer RNAs (eRNAs)

A new dimension to gene expression regulation has emerged from the discovery that active enhancers are broadly transcribed (reviewed in (Lai and Shiekhattar, 2014; Lam et al., 2014)). Although this type of transcription was originally identified at canonical enhancers, more recent high-throughput sequencing methods have demonstrated enhancer-derived transcripts (eRNA) to be widespread (Hah et al., 2013; Kim et al., 2010; Wang et al., 2011). In 2010, Greenberg and colleagues reported that enhancers activated by neuronal depolarization were associated with RNA Pol II and produced bi-directional transcripts. More importantly, they found that the levels of eRNA expression at these enhancers correlated with the production of messenger RNA (mRNA) at selected nearby genes (Kim et al., 2010). While the functional relevance of eRNAs is currently under debate, accruing evidence points toward an elaborate and indispensable role of eRNAs in both the construction of functional enhancers and the activation of gene expression. Several studies have shown that interference of eRNAs can lead to reduced transcriptional expression. For example, estrogen bound to its receptor in MCF-7 cells (breast adenocarcinoma) upregulates a set of estrogen-dependent coding genes by interacting with enhancers and inducing transcription of eRNAs (Li et al., 2013). However, elimination of the

associated eRNAs, using both RNA interference (RNAi) and antisense oligonucleotides (shRNA), reduced the ability of enhancers to positively influence gene expression. Interestingly, deletion of eRNAs in this system also reduced chromatin looping between enhancers and promoters. In the same study it was shown that eRNAs are also capable of upregulating transcription independent of the enhancer itself. Tethered eRNAs to synthetic promoters driving expression of luciferase reporter systems results in the increased expression of the reporter gene. Furthermore, a more recent study reported that in stimulated MEL cells (murine erythroleukemia), chromatin looping and eRNA expression precedes transcriptional activation of the target gene (Kim et al., 2015).

Remodeling proteins targeting nucleosome-associated histones dynamically translate environmental information into functional outputs. Specific combinations of histone modifications signal the recruitment of regulatory proteins and complexes that mediate the production of eRNA and together modulate gene expression. The high complexity of this process is underlined by the large number of histone modifying enzymes, multiple isoforms, and the many histone residues targeted for alterations. This system permits the execution of critical developmental programs, such as growth, differentiation, and normal cell death. Nonetheless, aberrant activity from some of the same remodeling enzymes often leads to tumor formation.

1.6 HDAC inhibitors

Carcinogenesis is a result of the accumulation of aberrant processes which lead to genetic instability and alterations in gene expression (Hanahan and Weinberg, 2011). Deregulation of HDACs can be a driver of tumorigenesis and other diseases. There are two prominent types of HDAC deregulation in cancer: a general over-accumulation of the metalloenzymes at the protein level and the expression of fusion proteins responsible for their inappropriate genomic

recruitment ((Grignani et al., 1998) and reviewed in (West and Johnstone, 2014)). In fact, it is estimated that up to 40% of human cancer tissues, including cancer of the colon, prostate, breast, lung, ovary, stomach, esophagus, pancreas, and thyroid, display relatively high expressions of most class I HDACs (1, 2, and 3) (Nakagawa et al., 2007). Furthermore, less differentiated and therefore more aggressive tumors are characterized by the strong overexpression of HDAC2 and HDAC3 (Müller et al., 2013). It is believed that the HDAC-induced hypoacetylated state contributes to transcriptional repression and maintains a positive environment for cell proliferation, survival, and anti-differentiation.

The aberrant HDAC activity found in many cancers makes these enzymes critical targets for therapeutic approaches. Consequently, HDAC inhibitors have emerged as promising therapeutics for the treatment of a wide variety of diseases, including cancers. Not surprisingly, these compounds have captured the interest from academic researchers and biotechnology entrepreneurs. Indeed, during the past decade a large number of HDACIs have been intensively investigated with the aims to introduce these compounds into the market.

1.6.1 Classifications of HDAC Inhibitors

Butyrate was the first compound reported with the ability to inhibit histone deacetylation in a variety of cultured cells (Candido et al., 1978). Over the past decade, multiple structurally diverse natural and synthetic HDAC inhibitors, with variable efficiency (nanomolar to millimolar concentrations), have been discovered and several have been subsequently improved through functional group modifications (Table 1.6). There are four main classes of HDAC inhibitors: (1) simple short-chain fatty acids, such as butyrate and valproic acid; (2) cyclic tetrapeptides including largazole and romidepsin (FK228); (3) hydroxamates such as vorinostat (SAHA), trichostatin A (TSA), belinostat (PXD101), and panobinostat (LBH-589); and (4) the benzamide

class of inhibitors which includes entinostat (MS-275), chidamide (CS055/HBI-8000), and mocetinostat (MGD0103).

Table 1.6 Molecular characteristics and target malignancies of HDAC inhibitors by class
HDAC (Histone deacetylase inhibitor); HDACI (HDAC Inhibitors); SAHA (Suberoylanilide hydroxamic acid); CTCL (Cutaneous T-cell lymphoma); PTCL (Peripheral T-cell lymphoma); N/A (not available); * (FDA approved). Adapted from (Bolden et al., 2006; Federico and Bagella, 2011; West and Johnstone, 2014; Xu et al., 2007).

HDACI class	Compound	HDAC specificity	Type of cancer
Short-chain fatty acids			
	Butyrate	Class I, IIa	Prostate cancer
	Valproic acid (VPA)	Class I, IIa	Colorectal, Prostate, Breast, Melanoma
	AN-9	N/A	Acute leukemia
Cyclic tetrapeptides			
	Largazole	Class I, IIb	Colorectal, Prostate, Ovarian, Melanoma, Hepatoma
	*Romidepsin (FK228)	Class I	CTCL, PTCL
Hydroxamates			
	*Vorinostat (SAHA)	Class I, II	CTCL
	Trichostatin A (TSA)	Class I, II	Breast carcinoma
	*Panobinostat (LBH-589)	Class I, II	Multiple myeloma
	*Belinostat (PXD101)	Class I, II	PTCL
	Givinostat (ITF2357)	Class I, II	JAK2 ^{V617F} -expressing myeloproliferative neoplasm
	Quisinostat (JNJ-26481585)	N/A	CTCL
	Pracinostat (SB939)	Class I, II, and IV	Prostate cancer and myelofibrosis
	Chidamide (CS055/HBI-8000)	N/A	Solid tumor and lymphomas
	Dacinostat (LAQ824)	Class I, II	Multiple myeloma
	Pyroxamide	Class I, unknown effect on class II	Neuroblastoma, prostate, and bladder carcinoma
	Abexinostat (PCI-24781)	Class I, II	Advanced and metastatic solid tumors
Benzamides			
	Entinostat (MS-275)	HDACs 1, 2, 3, 8 (marginally)	Melanoma
	Chidamide (CS055/HBI-8000)	N/A	Solid tumor and lymphomas
	Mocetinostat (MGD0103)	Class I and IV	B cell malignancies
	CI-994 (tacedinaline)	N/A	Acute myelocytic leukemia

Synthetic hydroxamates inhibit zinc-dependent HDACs by reversibly binding the metal within the target enzymes. Hydroxamates preferentially target class I, class IIa and IIb deacetylases. These types of inhibitors display promising activity in both *in vitro* and *in vivo* disease models. Within this type of HDACIs, there are currently numerous examples under clinical development and three have reached FDA approval (Figure 1.4). Vorinostat (marketed by Merck as Zolinza) was the first HDAC inhibitor to receive federal approval and it is used to treat patients with recurrent cutaneous T-cell lymphoma (CTCL) (Mann et al., 2007). Similarly, belinostat (Beleodaq) is currently used for the treatment of peripheral T-cell lymphoma (PTCL)

and just last year the latest hydroxamate panobinostat (Farydax) was approved for the treatment of multiple myeloma (MM).

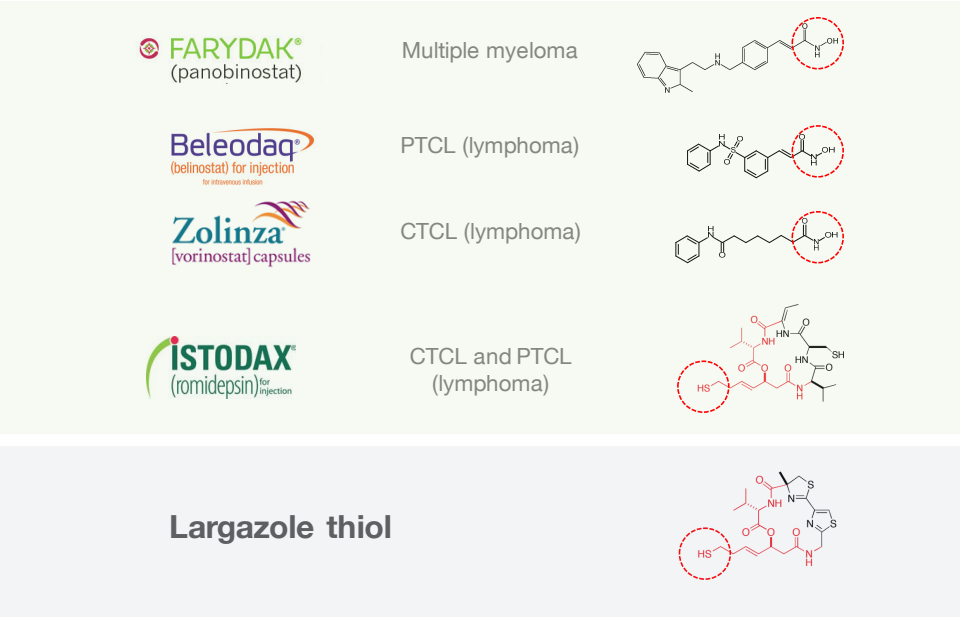


Figure 1.4 Largazole thiol and HDAC inhibitors approved by the US Food and Drug Administration (FDA) Shown are the structures of the hydroxamic acid based HDAC inhibitors approved by the FDA and the targeted malignancy: panobinostat (Farydax), belinostat (Beleodaq), and vorinostat (Zolinza). Romidepsin (Istodax) and largazole thiol belong to the cyclic tetrapeptide class of HDAC inhibitors. The zinc binding domain for each inhibitor is denoted with a red dotted circle. Adapted from (Cole et al., 2011; Mottamal et al., 2015).

Romidepsin (Istodax) is currently the only cyclic tetrapeptide on the market and is primarily used for the treatment of both CTCL and PTCL. Cyclic tetrapeptides, including largazole, are structurally different when compared to the less elaborate frames of hydroxamates (Figure 1.4). Most often found as natural products, cyclic tetrapeptides are prodrugs that require intracellular disulfide reduction in order to liberate the zinc-binding thiol group. One key structural feature of these inhibitors lies in a characteristic macrocyclic skeleton that provides additional binding interactions with the target enzymes. It has been suggested that this feature likely contributes to higher potency, stability, and selectivity towards different HDAC isoforms.

1.6.2 HDAC inhibitors' general mode of action

Inhibitors of metal dependent HDACs were first discovered as inducers of cell growth arrest, differentiation, and apoptosis (Candido et al., 1978; Davie, 2003; Haumaitre et al., 2008; Marks et al., 2000; Svechnikova et al., 2008). Early studies also determined that the pro-apoptotic effects of HDAC inhibitors were selective for transformed over normal cells. Although the vast majority of studies have been conducted using HDACIs belonging to the hydroxamate class of inhibitors (i.e. SAHA and TSA), it is believed that in general most HDACIs function by targeting similar pathways.

Inhibitors of class I and II HDACs, such as SAHA, suppress the growth of prostate cancer cells *in vitro* and *in vivo* (Butler et al., 2000, 2002). While the exact mechanism underlying this growth inhibition have not been fully elucidated, it appears that a standard biochemical hallmark of HDACIs is the transcriptional repression of cell cycle associated genes, including p16, cyclin D1, and Cdk6 and the reactivation of p21 (Finzer et al., 2001; Gui et al., 2004; Mathew et al., 2010; Zupkovitz et al., 2010). Correspondingly, the majority of these G1-associated gene transcripts are frequently deregulated in many classes of transformed cells, including colorectal tumors (Arber et al., 1999). Nonetheless, previous studies from our group using a stable p21-knockdown prostate cancer DU145 cell line did not show significant differences in cell cycle arrest nor apoptotic events when compared to control cells under similar treatments with largazole (unpublished data).

It is well-established that stimulation of apoptosis by HDACIs is achieved through multiple pathways. Intrinsic apoptotic activation is believed to be mainly achieved through upregulation of pro-apoptotic factors (Bolden et al., 2013). HDACIs also increase the expression of cell-death receptors which leads to activation of the extrinsic apoptotic pathway (Ashkenazi,

2002). Oxidative stress has also been proposed as a potential apoptotic mechanism of action deployed by HDACIs (Falkenberg and Johnstone, 2014). Cell exposure to HDACIs often results in the increase of reactive oxygen species and subsequent mitochondrial damage (Xu et al., 2006).

In general, HDACIs induce drastic changes in gene expression from cancer and normal cells and studies looking into the transcriptional effects from these compounds often report about 1,000 significant protein coding mRNA changes (Greer et al., 2015; Rafehi et al., 2014; Wang et al., 2009b). Changes in transcript levels induced by HDACIs are believed to be the major contributor to their therapeutic benefits and the degree of transcriptome alteration is likely dependent on the targeting and inactivation of specific HDAC isoforms and non-histone proteins.

1.6.3 Largazole

Largazole is a relatively novel natural anti-proliferative HDACI with unprecedented potency and high selectivity for cancer over normal cells (Taori et al., 2008). Largazole is unique in that it preferentially targets class I (HDAC1, 2, 3) and class IIb (HDAC10) HDACs. In vitro analyses demonstrate its unprecedented efficacy with sub-nanomolar inhibitory concentrations (Hong and Luesch, 2012). The relatively few number of steps required for the total synthesis of largazole and largazole-analogs combined with a high overall yield has driven extensive investigations on structural requirements for selectivity and antiproliferative properties (Bowers et al., 2009; Cole et al., 2011; Ying et al., 2008).

Largazole is a prodrug whose structure is composed of a macrocyclic ring stabilized by amide and ester linkages (depsipeptide) (Figure 1.5A). Attached to the cyclic structure is a lipophilic thioester side chain responsible for the enzymatic-inactivation and antiproliferative effects (Zeng et al., 2010). Upon protein-assisted hydrolysis of the thioester, the liberated thiol

group can coordinate the catalytic Zn^{2+} ion of the metal dependent HDAC enzymes (Cole et al., 2011).

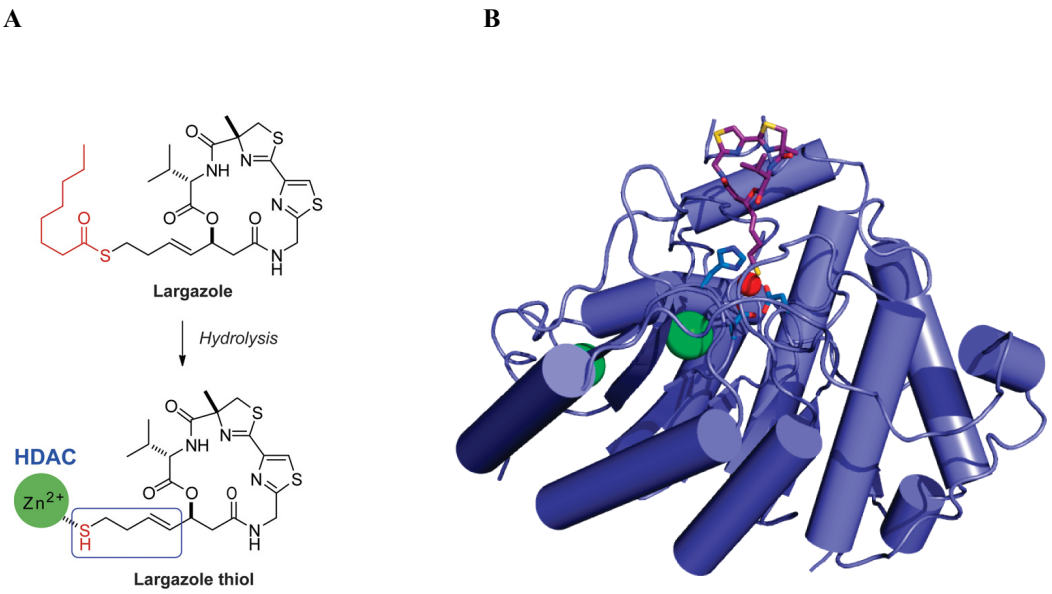


Figure 1.5 Largazole requires hydrolytic activation to complex with HDACs
A) Activation of largazole by protein-mediated hydrolysis generates largazole thiol. Liberation of the thiol group (red) permits the interaction with the HDAC catalytic Zn^{2+} ion (green). B) Crystal structure of HDAC8–largazole thiol complex at 2.14 angstrom-resolution. The catalytic Zn^{2+} ion (red sphere) on the enzyme is coordinated by a histidine and two aspartic acids residues (blue sticks). Largazole thiol is illustrated as a stick frame (C = magenta, N = blue, O = red, and S = yellow). Adapted from (Cole et al., 2011; Liu et al., 2010).

The crystal structure of the largazole-HDAC8 complex revealed that the negatively charged thiol of the inhibitor extends deep into the active site of the enzyme where it coordinates the catalytic Zn^{2+} ion (Figure 1.5B). The resulting thiolate-zinc coordination geometry is almost perfectly tetrahedral and is presumably the main contributing factor to the exceptional enzyme-inhibitor affinity. The complex is further stabilized by the formation of hydrogen bonds and polar and van der Waals interactions between the largazole thiol atom and numerous active site residues in the enzyme. Noticeably, the cyclic structure of largazole experiences minimal conformational

modifications once bound to HDAC8, whereas the enzyme undergoes multiple structural changes to accommodate the interaction (Cole et al., 2011).

At least two independent biological characterizations, including our own, have shown that largazole displays excellent selectivity at low nanomolar concentrations for colon cancer, leukemia, non-small cell lung cancer, CNS cancer, melanoma, ovarian cancer, prostate cancer, as well as breast cancer cell lines (Liu et al., 2010). The growth inhibitory effect largazole has been established to be based on its ability to inhibit classical deacetylase enzymes (Bowers et al., 2008; Zeng et al., 2010).

Thus far we know that largazole can affect the cell cycle through various mechanisms and this is clearly delineated by the distinct cell cycle phase arrests observed in multi-dose cell cycle analysis. Relatively low concentrations (~4 nM) prevent a G1 to S transition, while mid concentrations (~40 nM) cause a modest G2/M cell cycle arrest (Liu et al., 2010). It has been suggested that these differences are most likely indicative of target-dependent dynamics or as the result of secondary/downstream effects, but the corresponding molecular players in largazole-induced cell cycle arrest are yet to be identified (Liu et al., 2010). Another proposed mechanism by which largazole might be exerting its potent antiproliferative activity is through the downregulation of several cancer-associated growth factor receptors, including EGFR, HER-2, and MET (Liu et al., 2010).

Largazole can also inhibit cytokine-dependent activation and induces apoptosis of hepatic stellate cells (HSCs). The apoptotic response observed in HSCs was attributed to transcriptional downregulation of both anti-apoptotic genes bcl-2 (BCL2) and bcl-xL (BCL2L1); however no further validation was provided (Liu et al., 2013b). Importantly, largazole showed no effect on the proliferation of the immortalized human hepatocyte cell line MIHA, suggesting that the anti-

proliferative and apoptotic effects of the HDACI might be highly dependent on the proteome and/or epigenome context of targeted cells.

Largazole, as well as its ester and ketone analogs, selectively inhibits ubiquitylation of $p27^{Kip1}$ and TRF1 *in vitro* by blocking ubiquitin activation at the adenylation step (Ungermannova et al., 2012). The restrictive effect of largazole towards the E1 ligase is independent of its inhibitory activity towards HDACs. The thioester group is not required for inhibition; instead the macrocycle core and aliphatic tail were shown to play an indispensable role in preventing the activation of E1 ligase. The multiple inhibitory routes that largazole displays towards transformed cells makes the elucidation of primary factors driving apoptosis a difficult endeavor. However, it is also this feature of largazole that circumvents the problem of drug resistance caused by mutations or through cross-talk between signaling pathways.

HDAC inhibitors have emerged as promising compounds for the treatment of cancers. Although there are only four HDACIs that have been approved by the FDA for the treatment of lymphomas and multiple myeloma, there are over a dozen compounds that are currently under intense clinical investigations destined to be used as single agents or in combination therapies (Federico and Bagella, 2011; West and Johnstone, 2014). Further improvement of specificity and potency from these anticancer agents require both the identification of structural features utilized for targeting specific substrates as well as a deeper understanding of the epigenetic deregulations responsible for tumorigenesis. The ultimate scenario in regards to clinical treatment involves patient selection criteria based on specific biomarkers paired to the most adequate drug; consequently, the link between the biological function and the therapeutic benefit of HDACIs needs to be further investigated. largazole is among the most potent HDAC

inhibitors found yet, displaying similar anticancer properties as the FDA approved romidepsin but the underlying selectivity for cancer cell apoptosis remains poorly understood.

Chapter 2: Dose-dependent inhibition of histone deacetylases reprograms gene expression through global remodeling of the enhancer landscape

2.1 Introduction

Reversible lysine acetylation is important for homeostatic regulation of many cellular processes. The best characterized proteins that are subjected to this mode of regulation are histones. Lysine residues in the N-terminal tail of histone proteins are subjected to acetylation and deacetylation catalyzed by enzymes known as histone acetyltransferases (HATs) or histone deacetylases (HDACs) (Strahl and Allis, 2000). The patterns of histone acetylation dictate the epigenetic state of chromatin and transcriptional activity. Disruption of histone acetylation and deacetylation has been observed in human tumors. Indeed, transcription of tumor suppressor proteins is frequently silenced in tumor cells due the hyper- or aberrant activity of HDACs (Minucci and Pelicci, 2006). Accordingly, HDACIs are used clinically for the treatment of a subset of hematologic tumors (Marks, 2010).

There are 18 HDAC enzymes encoded by the human genome that belong into four distinct classes (Marks, 2010). Classes I, II and IV all contain a zinc (Zn^{2+}) ion in their catalytic site and are inhibited by pan-HDACIs such as Vorinostat, Belinostat, or Panobinostat (Marks, 2010). Class III comprises the mechanistically distinct NAD^{+}-dependent sirtuins. Collectively, HDACs are involved in regulation of diverse cellular functions including transcriptional regulation, DNA replication, DNA repair, cell signaling, migration, and differentiation (Haberland et al., 2009; Marks, 2010). It has been long recognized that HDACs are predominantly involved in transcriptional repression as loss of histone lysine acetylation, a hallmark of transcriptionally active open chromatin, decreases chromatin access. HDACs often exist as the catalytic module of chromatin remodeling machineries, including CoREST, NuRD,

Sin3, and N-CoR (Yang and Seto, 2008). These molecular complexes contain distinct HDACs and target specific genomic regions through sequence-specific interactions with non-histone proteins such as transcription factors, methyl binding proteins (MBDs), or other epigenetic modifier enzymes such as DNA and histone methyltransferases (DNMTs or HMTs). Aberrant recruitment of HDACs seen in cells with chromosomal translocations or mutations in certain transcription factors contribute to development of tumors (Minucci and Pelicci, 2006). Hence, HDACIs are used to de-repress the silenced genes for cancer treatment (Marks, 2010). As expected, transcriptome analysis in the presence of HDACIs revealed the drastic up-regulation of numerous genes (Kim et al., 2013; Rada-Iglesias et al., 2007; Rafehi et al., 2014). The same analysis also yielded a surprising finding that transcription of many genes is also repressed by HDACI exposure (Kim et al., 2013; Rada-Iglesias et al., 2007; Rafehi et al., 2014). The importance of HDACIs as anticancer therapies warrants more in-depth understanding of their activities in transcriptional regulation.

The opposing functions of HDACIs on transcription are difficult to reconcile. Genome-wide HDAC localization analyses indicate that HDACs are associated strongly with actively transcribed genes in human cells (Wang et al., 2009b). Divergent activities of HDACIs on transcription could be a result of deacetylation activity towards different classes of targets. For example, non-histone substrates, including certain transcription factors, are activated when deacetylated (Chen et al., 1999; Wolf et al., 2002; Xu et al., 2003). A recent study suggests that HDACIs target the transcription elongation complex and cause redistribution of other elongation factors across the genome (Greer et al., 2015). A major drawback of many genome-wide studies with HDACIs is the use of pan-HDACIs at a single dose. Given that different classes of HDACs

are associated with distinct IC_{50} values for a specific HDACI, it is not known whether the changes seen with HDACI treatment are relevant to their biological activities.

All HDACIs, except Romidespin, on the market are pan-HDACIs (i.e. target all 11 HDAC enzymes at varied degrees of inhibition) (Lane and Chabner, 2009; Marks, 2010). Despite the success of treatment of cutaneous T-cell lymphoma (CTCL), peripheral T-cell lymphoma (PTCL), and multiple myeloma (MM), there is limited success of HDAC inhibition in solid tumors (Lane and Chabner, 2009; Marks, 2010). There is a lot of enthusiasm for the development of combinatorial therapies using HDACIs. However, pan-HDACIs have significant dose-limiting toxicity which hampers their use in combination with other drugs. Isoform-selective HDACIs are more attractive for combination therapies (Lane and Chabner, 2009). The therapeutic benefits of HDAC inhibition are thought to be associated with their chromatin remodeling activities and the resulting transcriptional reprogramming changes. However, the molecular mechanisms by which HDACs and their inhibition regulate gene expression are still not fully understood. Largazole, a marine natural product discovered in cyanobacteria, is a highly potent Class I, Class IIb, and Class IV selective HDACI and displays selective killing of tumor cells (Bowers et al., 2008; Taori et al., 2008). Largazole offers a unique tool to address the mechanism of HDAC inhibition in cancer biology due to its selectivity, superb potency, and minimal off-target activities. Parsing out various mechanisms underlying largazole-induced transcription activation and repression could offer fundamental mechanistic insights critical for developing superior HDACIs with better clinical efficacy and low toxicity.

Here we present a comprehensive analysis of the specificity and molecular mechanisms of action for largazole. We show that largazole selectively inhibits class I and class IIb HDAC enzymes at a subnanomolar range and causes cytostatic responses in a variety of tumor cell lines.

We performed genome-wide studies to identify histone marks and gene signatures whose dose-responsive changes, upon exposure to increasing concentrations of largazole, closely match the cellular growth inhibition of 50% (GI_{50}) curve of the cytostatic response using ChIP-seq and RNA-seq. Our data show largazole induces profound dose-dependent changes in H3K9ac, H3K27ac, H3K4me1, H3K4me2, and the association of RNA Pol II with enhancers, promoters, and gene bodies. Low doses of largazole exposure resulted mostly in the up-regulation of gene transcripts whereas mid to high doses primarily triggered a depletion of mRNA. The decrease in RNA accumulation can be attributed in part to increasing RNA Pol II pausing. Although a subset of gene bodies exhibits elevation and spreading of both H3K9ac and H3K27ac marks upon largazole exposure, the dose response behavior is independent of the transcriptional response. Instead, our results reveal largazole causes remodeling of numerous enhancer elements by modulating H3K27ac but not H3K9ac in a dose-dependent manner and uncover a novel role for HDACs to maintain the repressive state of poised transcriptional enhancer elements.

2.2 Results

2.2.1 Largazole is a potent and selective HDAC inhibitor that suppresses tumor cell proliferation

The inhibitory activity of largazole towards a selective group of HDAC enzymes has been reported in several studies. However, different groups use different sources of enzymes, which leads to significant variability in HDAC *in vitro* assays, so it is difficult to relate results from one study to those from another. To comprehensively compare the HDAC selectivity profile of largazole to that of well-established HDACIs, largazole-thiol, Trichostatin A, and SAHA were independently incubated with each of the 11 HDAC enzymes *in vitro*. In agreement

with previous reports, we found that largazole-thiol is highly potent towards class I HDAC enzymes (particularly HDAC1, 2, and 3), and class IIb HDACs (6 and 10), with minimal inhibition of class IIa enzymes. In contrast, SAHA, and Trichostatin A exhibit inhibitory activity against all HDAC enzymes (Table 2.1). From this analysis, it becomes apparent that none of the HDAC inhibitors specifically target a single class of HDAC enzymes but instead have distinct selectivity profiles towards various isoforms of HDACs. Largazole potently inhibits class I, class IIb, and to a lesser extent class IV HDACs but spares class IIa HDACs.

Table 2.1 Largazole selectivity profile towards different class of zinc-dependent HDACs
Concentration-inhibition profiles of Largazole-thiol, Trichostatin A (TSA), and Suberoylanilide hydroxamic acid (SAHA) were determined in vitro. IC_{50} values are shown in molar concentrations.

Class	Enzyme	Largazole-thiol [M]	Trichostatin A [M]	SAHA [M]
Class I	HDAC-1	1.97E-10	7.48E-09	3.23E-07
	HDAC-2	4.15E-10	1.31E-08	9.19E-07
	HDAC-3	1.86E-10	3.04E-08	9.02E-07
	HDAC-8	1.21E-07	2.32E-07	8.98E-07
Class IIa	HDAC-4		6.51E-06	4.83E-05
	HDAC-5		2.49E-06	2.00E-05
	HDAC-7		2.45E-06	6.78E-05
	HDAC-9		2.45E-06	9.09E-05
Class IIb	HDAC-6	1.33E-08	1.09E-09	1.59E-08
	HDAC-10	4.22E-11	1.75E-08	1.09E-07
Class IV	HDAC-11	2.04E-10	1.57E-08	4.81E-07

To determine growth inhibitory properties of largazole in human tumor cell lines, we submitted largazole to a NCI developmental therapeutic program (DTP) and determined growth-inhibitory effect, GI_{50}; cytostatic effect, TGI_{50}; and cytotoxic effect LC_{50} concentrations against all NCI 60 cell lines (data not shown). We found that colorectal cell lines are particularly

sensitive to largazole and because HCT116 cells have been consistently investigated by genome-wide sequencing analyses, we chose this cell line for all follow-up studies. Cell cycle analysis of HCT116 cells by flow cytometry revealed that largazole stimulation for 25 hours alters cell cycle progression and leads to significant apoptosis (Figure 2.1A).

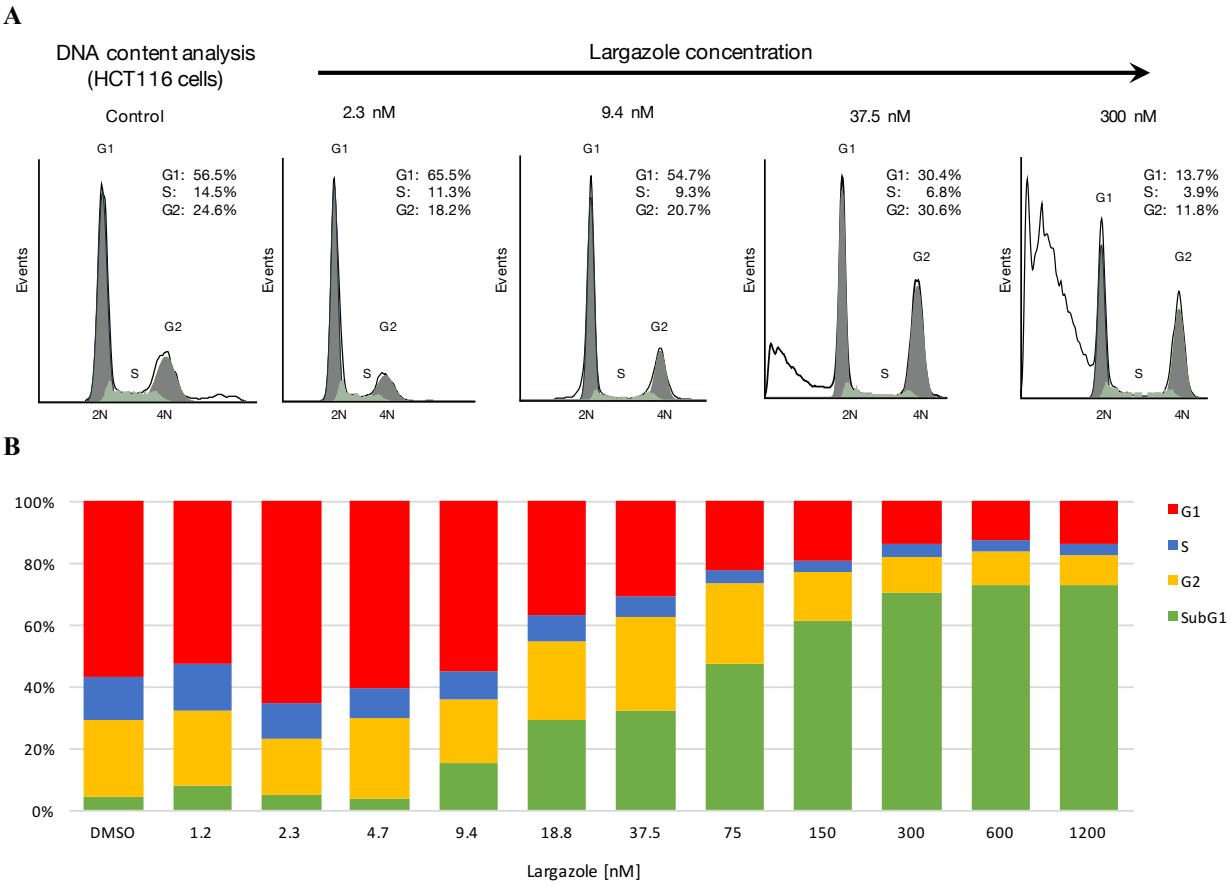


Figure 2.1 Effects of largazole on cell cycle progression and apoptosis in HCT116 cells
Quantitative analysis of the cell cycle progression by propidium iodide staining using flow cytometry in HCT116 cells treated with the indicated largazole concentration for 25 hours. Cell cycle distribution of propidium iodide (PI)-labeled cells was analyzed by flow cytometry. A) The peaks in the illustration correspond to G1, S, and G2 phases of the cell cycle. B) Histogram showing the percentages of cells in G1 (red), S (blue), and G2 (yellow) phases of the cell cycle as well as subG1 faction (green).

Specifically, exposure of HCT116 cells to low concentrations of largazole induces cell cycle arrest at the G1 (~4 nM) and G2/M (~37 nM) phases; however, these effects systematically dissipate at higher largazole concentrations (Figure 2.1). Significantly, largazole caused a

dramatic increase of sub-G1 phase (apoptotic) cells in a dose-dependent manner that becomes apparent at ~18.8 nM and plateaus ~300 nM (Figure 2.1B green). These results confirm that largazole inhibits proliferation and induces significant cell death of HCT116 cells at low nanomolar concentrations (GI_{50} = ~34 nM).

2.2.2 Largazole induces dose-dependent acetylation of H3K9, H3K27 and mono-methylation of H3K4

HDACIs alter the acetylation state of histone and non-histone proteins, and this effect is often studied with the use of pan-acetyl antibodies that recognize multiple acetylated lysine residues within a single histone. To determine largazole's effects on histone acetylation in mammalian cells, we analyzed nuclear extracts from HCT116 cells treated with increasing concentrations. Dose-dependent accumulation of acetylated lysine 9 and 27 on histone H3 (HK9/27ac) was observed with a significant increase in signal from cell extracts treated with ~18 nM (Figure 2.2).

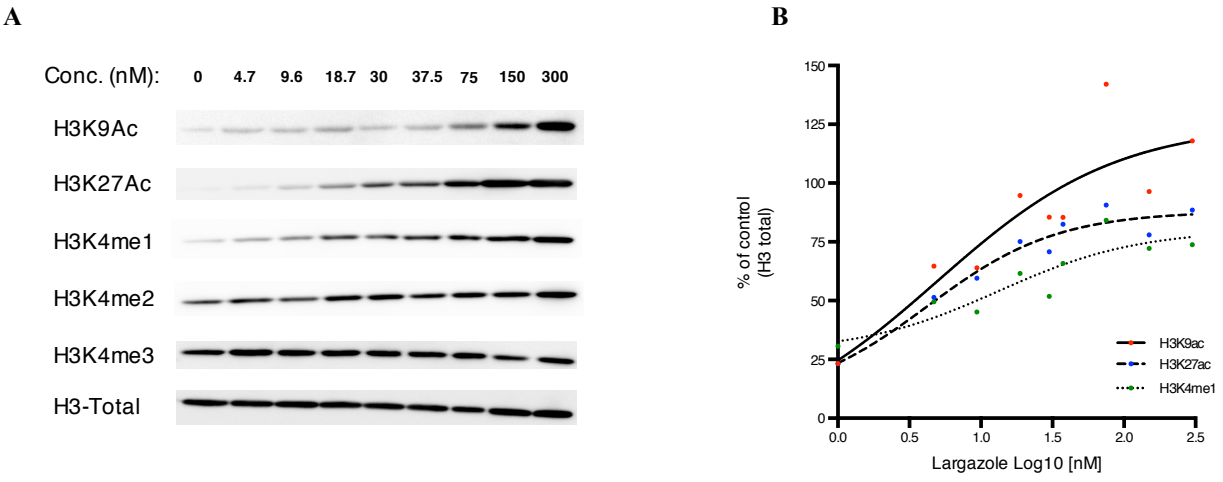


Figure 2.2 Largazole induces systematic acetylation of H3K9 and H3K27 and elevates the methylation of H3K4me1
A) Largazole-induced modifications of specific lysine residues of histone H3. Largazole increases H3K9 and H3K27 histone acetylation and mono-methylation of histone H3K4. Total histone H3 was used as a loading control. B) Dose dependent increase in H3K9ac, H3K27ac and H3K4me1 in largazole stimulated HCT116 cells. Quantified signal of histone acetylation and histone methylation were normalized to total histone H3.

To determine if largazole specifically affects histone acetylation, we also measured the three different methylation states of histone H3K4 in response to increasing largazole treatments. Interestingly, we observed that mono-methylation of H3K4 increased systematically with largazole dose treatments while di- and tri-methylation did not show significant changes. Based on these observations, we conclude that largazole elevates acetylation of histone H3K9, –K27, and H4Kme1 in a dose-dependent manner.

2.2.3 Genome-wide dose-dependent changes in acetylation of H3K9, H3K27 and mono-methylation of H3K4 in response to largazole treatment

Overexpression of HDACs has been observed in transformed cells and is thought to contribute to the repressive chromatin state harboring tumor suppressor genes. To investigate genome-wide acetylation and methylation changes conferred by largazole, we employed chromatin immunoprecipitation coupled with massive parallel sequencing (ChIP-seq) using antibodies targeting H3K9ac, H3K27ac, H3K4me1 and H3K4me2 in HCT116 cells treated with increasing concentrations of largazole.

Because largazole confers an extensive degree of newly acetylated genomic regions that can extend for several kilobases, traditional peak-calling algorithms were not suitable for analysis of H3K9ac and –K27ac ChIP-seq data (Figure 2.3A top). Instead, we used Fast Read Stitcher (FStitch) under default signal threshold to identify broad regions of enrichment over a wide range of signal strength (Azofeifa et al., 2014). The rest of the data was analyzed using MACS2 peak caller software with a $p = 1e^{-5}$ cutoff (Zhang et al., 2008). FStitch was originally designed for the detection of nascent RNA transcripts but the extensive similarities between global nuclear run-on sequencing (GRO-seq) and ChIP-seq data from largazole-induced hyperacetylation prompted us to investigate its application for the detection of extra-long range

signal coverage. We first tested FStitch efficacy by comparing the signal obtained in ChIP-seq data from untreated HCT116 cells (DMSO) to that of MACS2 peak calls.

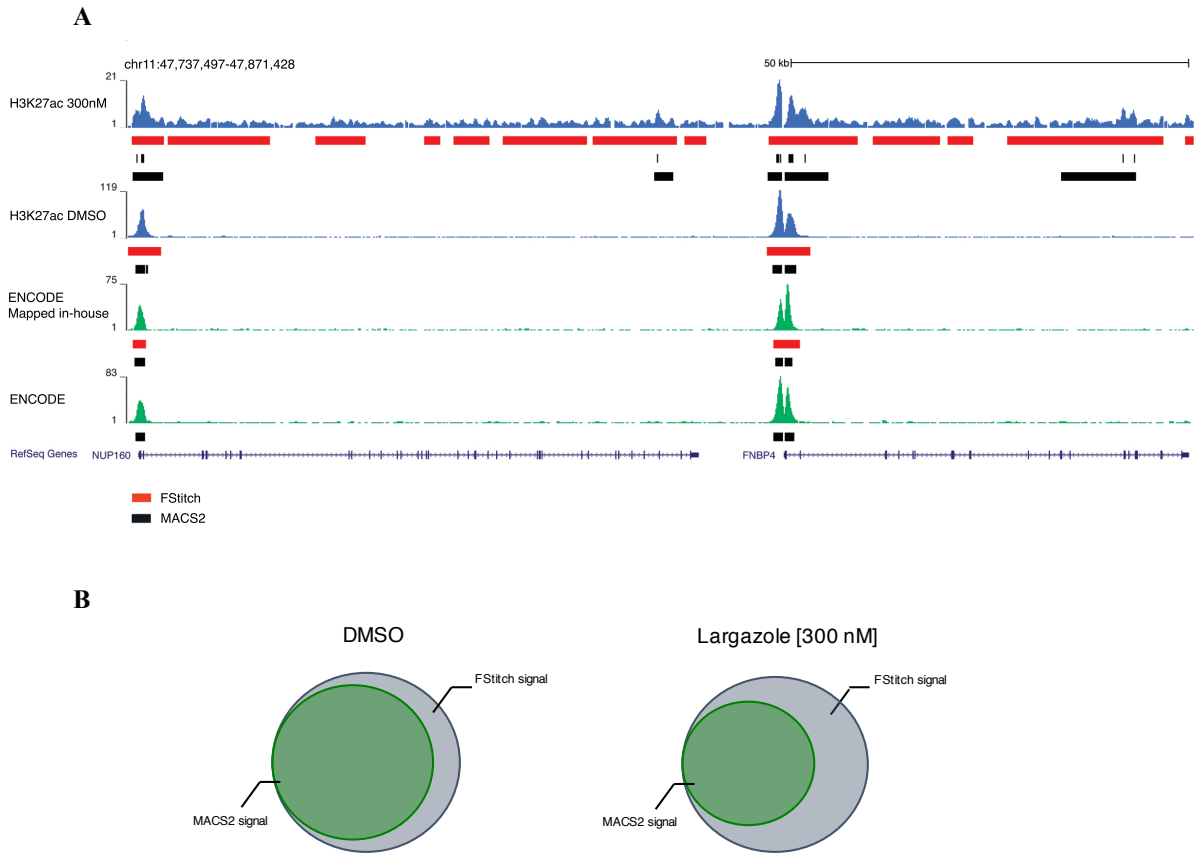


Figure 2.3 Comparison of two ChIP-seq signal calling algorithms (FStitch and MACS2)
A) A screen shot from Genome Browser (UCSC) showing ENCODE's H3K27ac (green) and in-house prepared and processed H3K27ac (blue) ChIP-seq data in HCT116 cells. The 125 kb genomic window illustrates the statistical significant regions called by FStitch (red) and MACS2 (black) using their default signal thresholds. B) Overlap between peaks called by each algorithm. Venn diagram shows ~26% of the peaks called by FStitch (blue) using domestic DMSO data were not detected by MACS2 (green) and ~50% when a similar comparison was performed to domestic 300 nM data.

FStitch showed superior detection range for well-defined acetylation peaks found in cells in the basal state; about 26% of the peaks called by FStitch using untreated ChIP-seq data were not detected by MACS2 (Figure 2.3B left). We then compared the two signal-calling algorithms using ChIP-seq data from HCT116 cells treated with 300 nM largazole. We found that MACS2 failed to detect ~50% of acetylation signal called by FStitch in the treated cells (Figure 2.3B right). This is not surprising, since the read density of hyper-acetylated ChIP-seq data is predominantly widely spread and low in profile. Therefore, FStitch was used to compare acetylation changes between HCT116 cells treated with DMSO and those treated with increasing doses of largazole (4.7 nM – 300 nM).

To gain a better understanding of the signal distribution of H3K9/27ac, we divided the human genome into five features: proximal promoter (+/- 2kb from the TSS), gene regions (-2kb from TSS to the end of the annotated gene), 3' end (end of annotated gene to 2kb downstream), intergenic, and enhancer elements (+/- 2kb from the determined center). These regions were annotated using the RefSeq hg18 gene reference from UCSC Genome Browser (Kent et al., 2002) and expressed as $\log_2$ fold ratios of the observed signal from DMSO-treated cells to that of largazole-treated HCT116 cells. In unstimulated cells, we detected ~5% of the genome associated with H3K9ac and K27ac signal with both marks heavily enriched at enhancer elements (54.0% for H3K27ac and 42.4% for H3K9ac) and transcription start sites (23.6% for H3K27ac and 26.1% for H3K9ac) (Figure 2.4A and 2.4B).

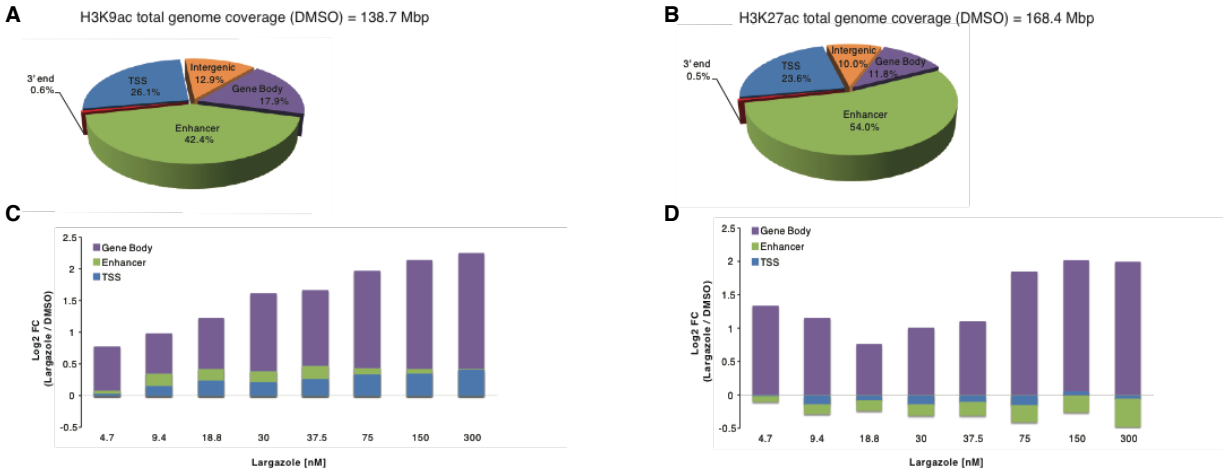


Figure 2.4 Largazole-induced enrichment of H3K9ac and H3K27ac organized by genomic features
A) Pie charts illustrate the distribution of H3K9ac and B) H3K27ac signals (as determined by FStitch) resulting from ChIP-seq experiments performed in vehicle (DMSO) treated HCT116 cells. H3K9ac and –K27ac signals are found more prevalent in enhancer regions (green) and TSS (blue), to a lesser extent in gene bodies (purple) and intergenic locations (orange) and rarely found in 3' ends (red). C) The log2 fold change ratio for increasing doses of largazole (nM) shows a general trend with gradual H3K9ac enrichment for all genomic features. In contrast, D) H3K27ac accumulation along enhancer regions diminishes with increasing doses of largazole.

Acetylation of histone H3 on lysine 9 increased at four of the genomic features analyzed: transcription start sites, 3' ends, enhancers, and more drastically along gene body regions (Figure 2.4C). In contrast, a clear reduction in histone H3K27ac was found along enhancer elements, and this loss was more prominent under higher largazole dose treatments (Figure 2.4D). In addition, both acetylation marks (H3K9ac and –K27ac) displayed a reduction along intergenic regions (Figure 2.5). The intergenic features contained classical narrow peaks for H3K9ac and –K27ac under basal cellular conditions. Because they were intergenic, we were unable to assign them into specific genomic categories; although it is possible that these regions contain a different class of *cis*-regulatory elements (e.g. lncRNAs) or un-annotated genes.

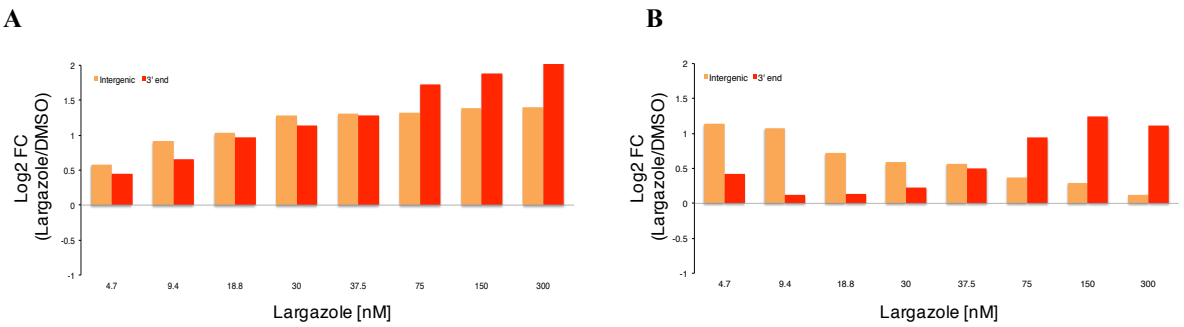


Figure 2.5 Largazole-induced enrichment of H3K9ac and H3K27ac along 3' ends and intergenic regions
A) The log2 fold change ratio for increasing doses of largazole (nM) shows a general trend with gradual H3K9ac enrichment for 3' ends and intergetic features, whereas, B) H3K27ac accumulation diminishes with increasing doses of largazole along genomic regions that are not associated with annotated features (intergenic).

Compared to H3K9/27, we found a small but significant increase in H3K4me1 ChIP-seq signal that is primarily localized to gene bodies and intergenic elements (data not shown). Largazole-induced mono-methylation of H3K4 at gene bodies appears to be the result of basal TSS signal migrating downstream, whereas intergenic regions often displayed the emergence of *de novo* H3K4me1 peaks, possibly due to newly established sites of enhancer deployment that lack detectable acetylation levels. Collectively, our data indicate that H3K9ac and H3K27ac undergo different dose-dependent changes in response to largazole, where loss of H3K27ac appears to be a prevalent feature at enhancer elements.

2.2.4 Distinct patterns of dose-dependent changes in acetylation of H3K9 and –K27 at actively transcribed genes

Since H3K9ac and –K27ac are functionally associated with proximal promoters and the most drastic acetylation changes occurred at protein-coding regions, we further examined genes that were occupied by H3K9ac and H3K27ac under basal conditions. A total of 10,356 unique genes with a minimal length of 3 kb were found bound by H3K9ac at the corresponding transcription start sites (TSS), and a similar screen yielded 10,272 genes positive for the

H3K27ac signal. As shown in Figure 2.6A, we found distinct gene patterns associated with H3K9 and –K27 acetylation changes. A particular class of genes displayed a TSS highly occupied by acetylated H3K27 or –K9, where the signal remarkably spread in a dose-dependent manner into the transcribable region (Figure 2.6A right panel). The second class of genes displayed a moderate amount of histone acetylation at the TSS and the signal remained anchored to the promoter throughout all largazole doses (Figure 2.6A left panel). Last, we found over 7,600 genes that were not associated with H3K9/27ac and remained acetylation free under all doses of largazole stimulation (Figure 2.6A middle panel). Using a cutoff of a two-fold increase in acetylation from untreated (DMSO) vs 300 nM largazole exposure, we found 4,155 unique genes exhibiting H3K27 hyperacetylation and 5,970 genes with higher levels of acetylated H3K9 (Figure 2.6B). The relative number of genes without significant acetylation (acetylation "deserted" genes) was very similar for both histone marks (Figure 2.6B). Thus largazole-induced dose-dependent changes in histone acetylation marks are both acetylation mark specific and restricted to defined genomic regions, suggesting separate regulation and functionality associated with these two marks.

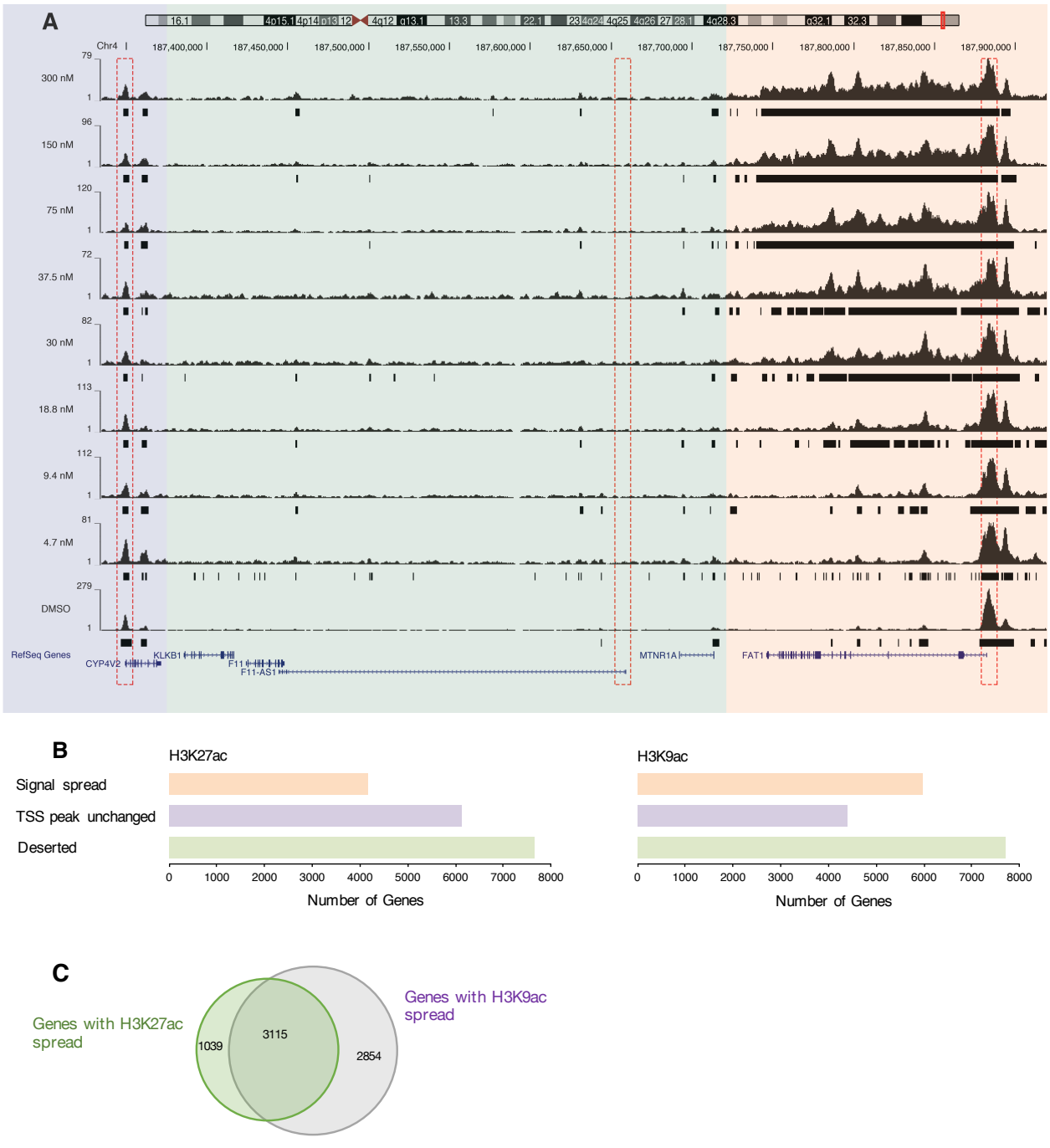


Figure 2.6 A selective set of gene regions display H3K9 and H3K27 hyperacetylation upon largazole treatment

A) A representative genomic snapshot of H3K27ac peaks illustrating different responses of gene bodies to newly acetylated histones. The signal initiates from the TSS (red dotted rectangle) of the FAT1 gene (pink panel) and spreads into the coding region or in the case of CYP4V2 (purple panel), the preexisting acetylated TSS remains unmodified throughout all largazole dose treatments. Genes that do not show H3K27 or –K9 acetylation at the TSS (green panel) under basal conditions do not associate with the two histone marks as a result of largazole treatment. B) Number of gene regions associated with the three response categories for H3K27ac (left) and H3K9ac (right). C) Venn diagram showing the number of genes that exhibit new association with H3K27ac (green), H3K9ac (purple), and those that display both acetylation marks.

2.2.5 Largazole induces dose-dependent changes in RNA transcripts independent of gene body acetylation patterns

Since H3K9ac and H3K27ac exhibit dose-dependent signal spreading, we wondered if these patterns are predictive of changes in gene expression. Largazole induces significant cell death only after 17 hours of exposure. Hence we harvested poly(A) RNA from HCT116 cells treated for 16 h with increasing concentrations of largazole and performed RNA-seq analysis. Only transcripts that exhibited dose-dependent up- or down-regulation by 1.5 fold were selected for further analysis. To examine transcriptional effects conferred by hyperacetylation of gene bodies, we plotted mRNA levels from all nine treatments for genes that showed significant acetylation spread. Surprisingly, genes hyperacetylated with H3K9ac or –K27ac showed up-regulation and down-regulation of transcription to similar extents, and a smaller set of genes had no significant changes in mRNA levels (Figure 2.7A and 2.7B). It is worth noting that a change in transcript levels from genes associated with H3K9ac occurs at lower largazole doses than for H3K27ac.

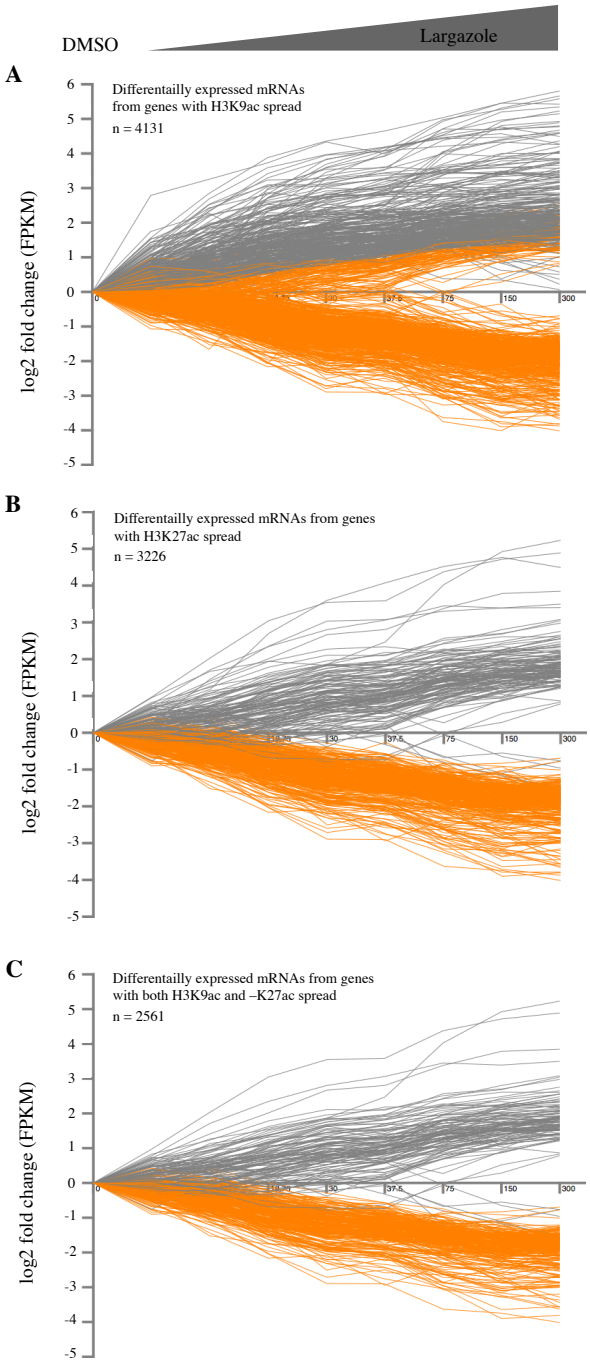


Figure 2.7 Hyperacetylation of gene regions does not predict higher mRNA accumulation
We used DREM version 2.0 software to visualize dynamic transcript changes as a function of largazole dose with a minimal absolute expression change of 1.5. Differentially expressed mRNAs from genes with only A) H3K9 hyperacetylation, B) H3K27 hyperacetylation, or C) those displaying an increase association with both histone marks.

To explore a possible synergy between H3K9ac and –K27ac with respect to transcriptional effects, we also looked at the mRNA expression levels of 3,116 genes that developed enrichment for both histone marks (Figure 2.6C). Similar to mRNA expression patterns from genes hyperacetylated at H3K9 or –K27, mRNAs from genes whose coding regions exhibited spread of both acetylation marks showed both events of up- and down-regulation (Figure 2.7C). Taken together, our findings show that elevation and spreading of H3K9ac or H3K27ac upon largazole exposure appears to be correlated with actively transcribed genes but fail to predict the direction of change (i.e., increase or decrease in transcript levels).

2.2.6 Largazole induces RNA Pol II pausing at a subset of genes

To examine the genome-wide effects of largazole-induced hyperacetylation on RNA Pol II occupancy along gene bodies, we conducted ChIP-seq experiments targeting total RNA Pol II in HCT116 cells treated with DMSO and those treated with 75 nM and 300 nM largazole. We used the 'pausing index' (PI) as the measurement to determine the extent of RNA Pol II pausing in a selected group of genes (n = 2352, refer to methods and materials for selection criteria). Pausing index was calculated by dividing the RNA Pol II ChIP-seq unique read density in the proximal promoter region by that in the gene body (Figure 2.8).

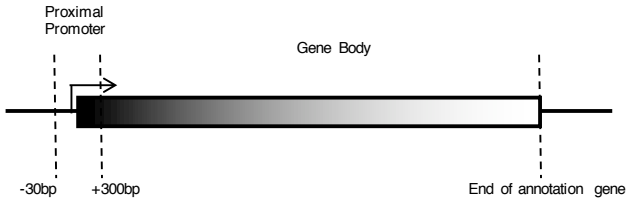


Figure 2.8 Parameters for pausing index analysis
Diagram shown illustrates the genomic locations used to calculate the pausing index (PI) for genes displaying a significant association with total RNA Pol II at the transcription start site, as determined by described criteria in materials and methods. Proximal promoter regions were assigned to a 330 bp window flanking the annotated start site (RefSeq Genes hg18), whereas the the remaining of the gene was declared as the gene body. The PI corresponds to the ratio of total RNA Pol II density (unique mapped reads) in the proximal promoter bin to that of the transcribed region. Adapted from (Rahl et al., 2010).

We observed that RNA Pol II PIs increased systematically upon largazole treatment for most of the ~2300 genes analyzed. As Figure 2.9A shows, most genes showed an increased PI when comparing 75 nM largazole-treated to DMSO-treated cells. Treatment with 300 nM largazole further increased PI relative to 75 nM treatment, but this increase was not as great as that between DMSO and 75 nM largazole-treated cells. Correlation comparisons of pausing indices from the three cellular conditions showed that increases in PIs are significant (refer to slopes in Figure 2.9B) and the magnitude and variability of such changes fall within a reasonable range ($R^2 = 0.722$ for DMSO vs 75 nM and $R^2 = 0.792$ for 75 nM vs 300 nM). For instance, the determined PI for transcription factor DP1 (TFDP1) in DMSO treated cells is 0.43 and the index increased to 3.70 in cells treated with 75 nM largazole. Depletion of RNA Pol II signal throughout the gene body is the main contributing factor to the increase of PI at the TFDP1 loci (Figure 2.9C). However, at the higher largazole dose treatment (300 nM), TFDP1 displays a PI of 16.36 that mainly reflects the vast accumulation of RNA Pol II restricted to the proximal promoter.

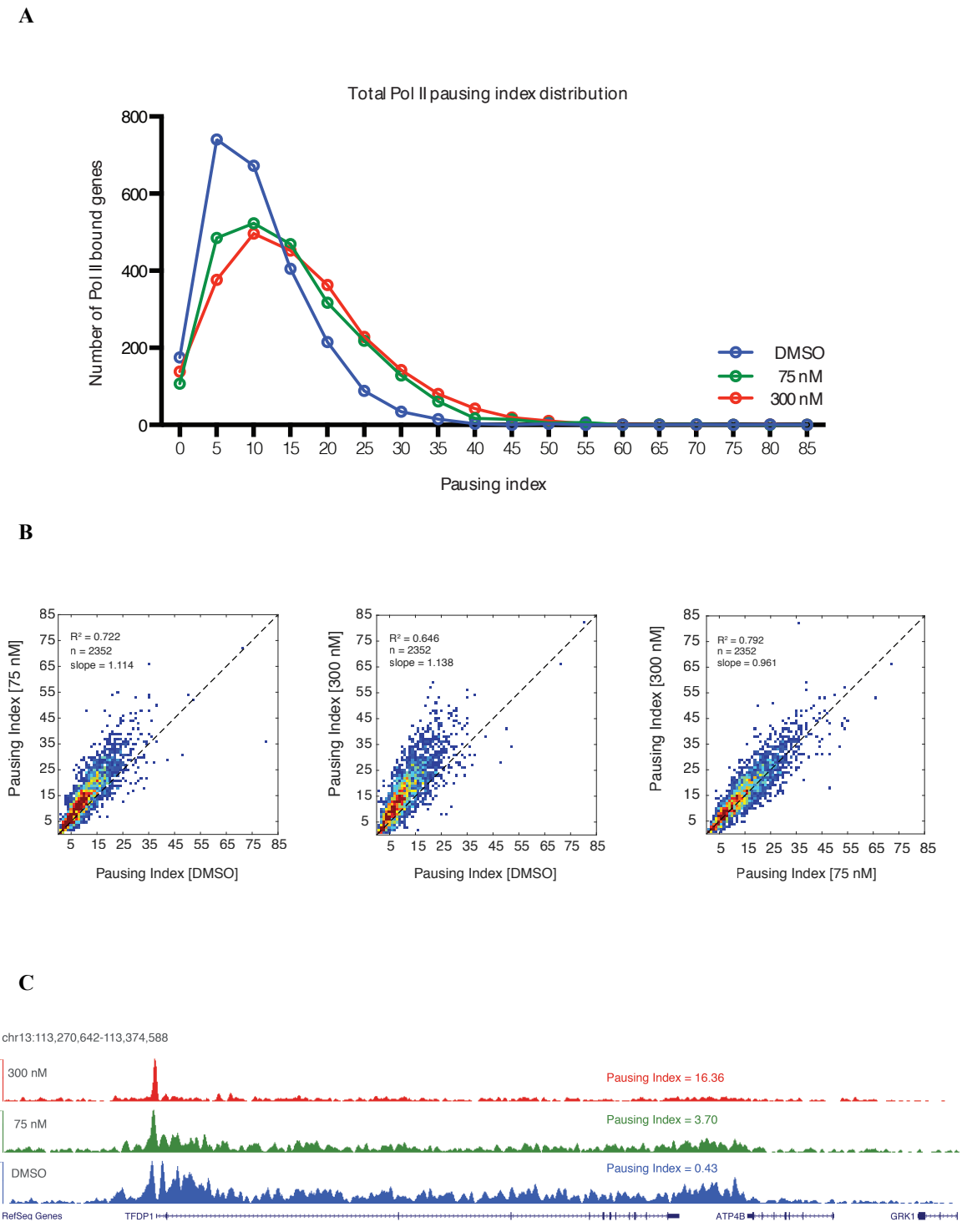


Figure 2.9 Effects of largazole on RNA Pol II pausing index
A) Histograms depicting the calculated PI distribution of a group of genes (refer to methods and materials for selection criteria) treated with DMSO (blue), 75 nM largazole (green), and 300 nM largazole (red). B) Contour plots showing Pearson correlation analysis between the calculated pausing indexes under the three different conditions. C) Screen shot of the TFDP1 loci showing total Pol II ChIP seq signal from HCT116 cells treated with vehicle (blue), largazole at 75 nM (green) and largazole at 300 nM (red) showing the calculated pausing index.

We then examined the influence of RNA Pol II pausing on the relative accumulation of transcripts associated with affected genes. In untreated cells, we found a modest but strongly supported (p = 2.2e-16) negative correlation (r = -0.333) between RNA Pol II pausing indices and relative accumulation of mRNAs (FPKMs) (Figure 2.10 bottom). We observed that highly expressed genes, such as TFDP1 and MYC (FPKMs > 80), were associated with relatively low PIs (0.46 and 2.74, respectively), whereas silent genes or those with low levels of expression (FPKMs < 1) such as BEST3 illustrated PIs greater than 20.

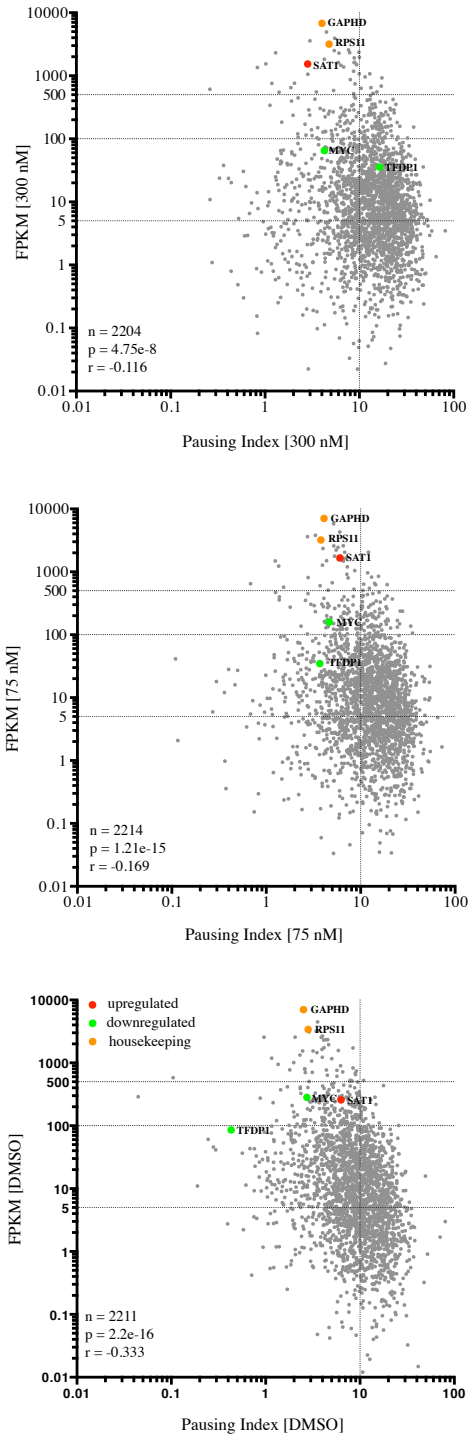


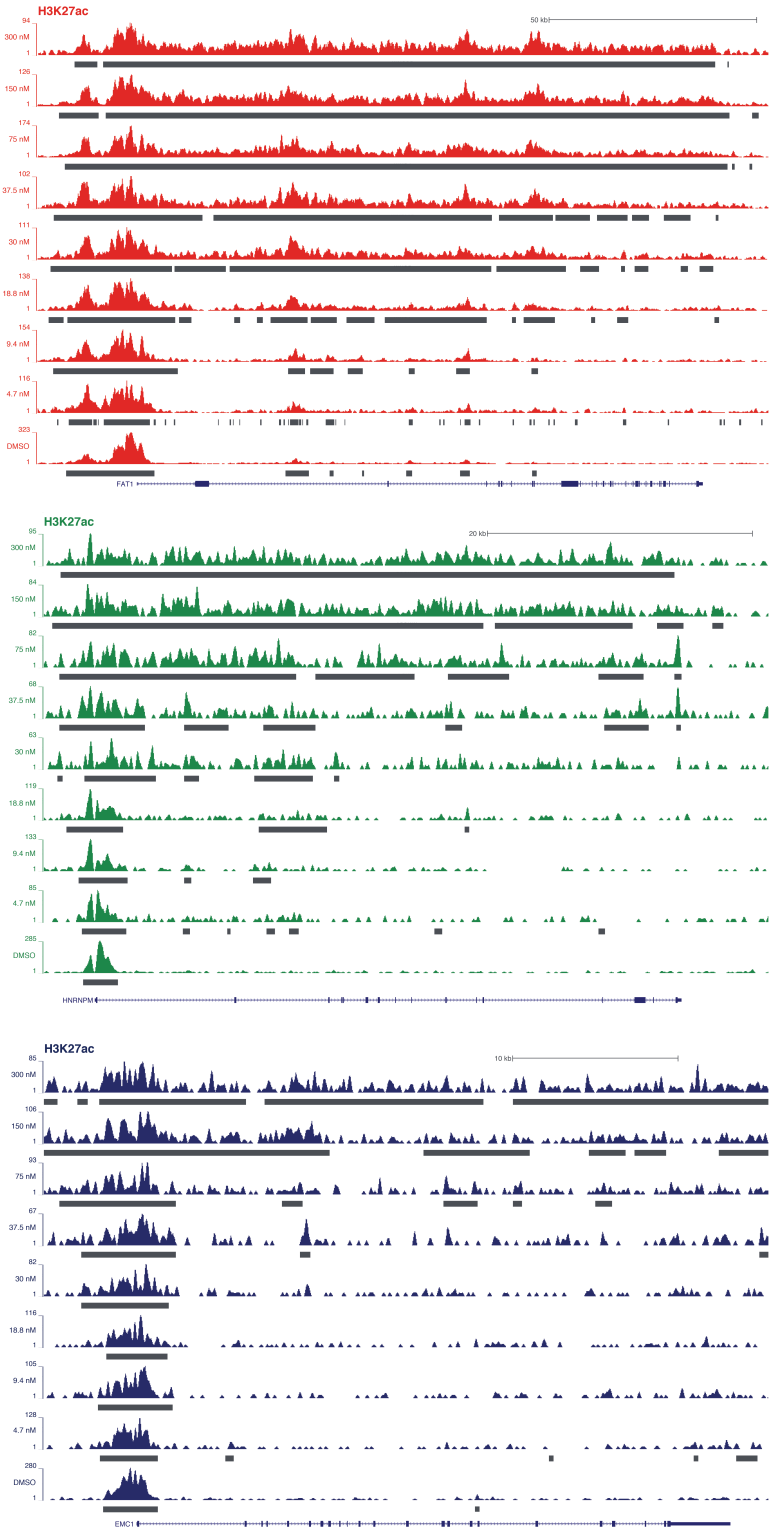
Figure 2.10 Correlation between pausing index and relative transcript levels (FPKMs)
Pearson Correlation method was implemented using the ln (PI) and ln (FPKM) values from ~2200 transcribable gene regions (grey). The input gene list was generated from gene bodies with a minimum length of 3 kb and that were bound by total Pol II at the transcription start site, as determined from MACS2 narrow peak calling signal. Overlapping gene regions and genes containing intergenic enhancers were excluded from the analysis. Three categories of transcriptionally regulated genes are shown as representative elements of the data; downregulated genes (green), upregulated genes (red) and housekeeping genes (orange).

Analyses of data from cells treated with 75 nM and 300 nM concentrations of largazole showed a marginal decrease in the correlation between PIs and FPKM values, however a general unidirectional trend of some genes became evident when looking at the changes in both PI and FPKM values from downregulated genes (Figure 2.10 middle and top). Increase in pausing indices is clearly the most dominant pattern observed among the analyzed gene regions; however, we also see a small number of genes with unchanged PIs and relatively constant FPKM values (GAPDH and RPS11) as well as genes that became less paused and transcriptionally upregulated (SAT1 and SIRT4). We conclude that for most genes analyzed, largazole specifically interrupts RNA Pol II occupancy downstream of proximal promoters and this event most likely affects the transition from initiation to elongation or the elongation steps of RNA synthesis.

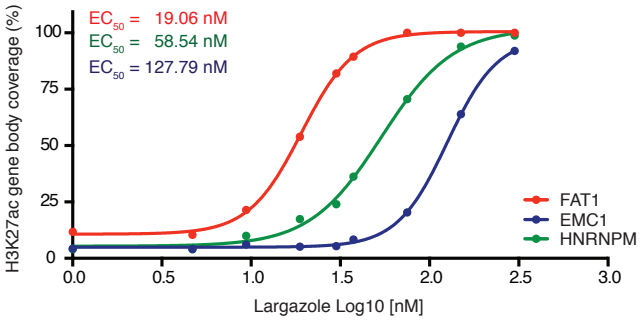
2.2.7 Low paused genes are more sensitive to H3K27 hyperacetylation

Class I HDACs preferentially occupy promoters of active genes and positively correlate with transcription levels (Wang et al., 2009b). Accordingly, it is expected that highly expressed (low paused) genes should be more sensitive to largazole and perturbations in the enzymatic activity of HDACs should tip the balance in favor of the HATs. This should be reflected in our data by the accumulation of newly acetylated H3K9 and –K27 signal along highly expressed genes in cells treated with relatively low concentrations of largazole. Indeed, among the genes displaying hyperacetylation of histone H3K9 and –K27, we noticed clear differences in dose-specific signal spread. For instance, the transcribable region of the protocadherin gene FAT1 displays a substantial association with H3K27ac in cells treated with 18.75 nM largazole, and the signal reaches complete gene body coverage in cells stimulated with 75 nM (Figure 2.11A top).

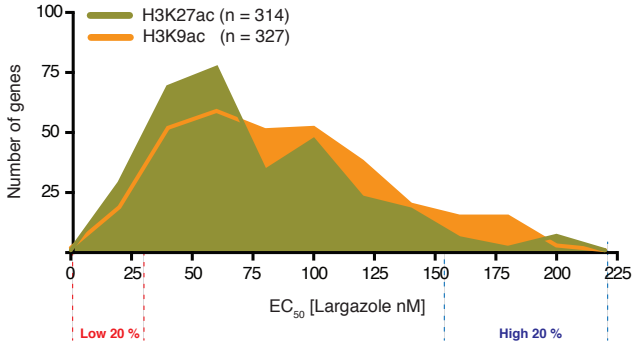
A


H3K27ac
50 kb
300 nM
150 nM
75 nM
37.5 nM
30 nM
18.8 nM
9.4 nM
4.7 nM
DMSO
FAT1
H3K27ac
20 kb
300 nM
150 nM
75 nM
37.5 nM
30 nM
18.8 nM
9.4 nM
4.7 nM
DMSO
HNRNPM
H3K27ac
10 kb
300 nM
150 nM
75 nM
37.5 nM
30 nM
18.8 nM
9.4 nM
4.7 nM
DMSO
EMC1

B



C



D

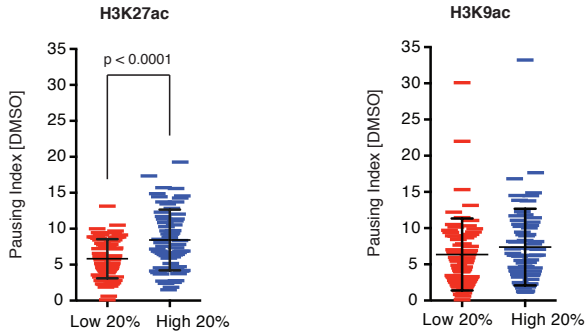


Figure 2.11 Correlation between histone acetylation signal spread sensitivity along gene bodies and pausing index

A) Snap shots from UCSC Genome Browser showing H3K27ac ChIP-seq signal over three different gene regions illustrate different dose-dependent acetylation of H3K27; low dose (red), mid dose (green), and high dose responders (blue). B) EC_{50} values calculated by dose response plots of the normalized H3K27ac gene body coverage (FStitch signal) for the three genes in A. C) Histograms showing the distribution of EC_{50} values, for both H3K27ac (green) and –K9ac (orange), associated with a set of selected gene regions. D) Student's unpaired t-test analysis between the calculated Pol II pausing indexes from genes with the lowest (20%) and highest (20%) EC_{50} values determined from the histogram shown in C. Individual genes with associated PI are plotted and the mean and standard deviation are shown with black horizontal lines.

We also found gene regions, such as the locus for HNRNPM, that do not associate with significant H3K27ac or –K9ac levels until ~30 nM largazole treatment (Figure 2.11A middle). Finally, there is a group of genes that are more resistant to hyperacetylation changes. For example, in the EMC1 gene region newly acetylated histones are only detected in cells treated with largazole concentrations at or above 75 nM (Figure 2.11A bottom).

To elucidate the sensitivity of each gene to largazole-induced acetylation changes, we determined the largazole concentration necessary to induce a half-maximal acetylation response (EC_{50}) in genes displaying 50% or greater H3K9ac or –K27ac signal coverage over the annotated gene lengths. Using FStitch calls, we calculated the total acetylation signal for both H3K9ac and –K27ac along the gene bodies for each of the nine corresponding ChIP-seq experiments and used the Sigmoidal Dose Response Search algorithm (SDRS) with a $p = 0.05$ cutoff (Ji et al., 2011). This approach allowed us to effectively assign each gene with a largazole dose sensitivity for both H3K9ac and –K27ac changes (Figure 2.11B). Distribution analysis of EC_{50} values from the two histone marks revealed a similar range of largazole dose sensitivity (~8 nM to ~210 nM). However, within the population of gene regions associated with H3K27ac, we observed a higher number of genes responsive to low concentrations of largazole as compared to H3K9ac (Figure 2.11C).

To explore a possible correlation between the basal pause state of genes and sensitivity to largazole-induced acetylation changes, we compared the PIs from the 20% of genes most sensitive to largazole (lowest EC_{50}s) to that of the 20% of genes exhibiting the most resistance to acetylation changes (highest EC_{50}s) (Figure 2.11C). We performed this analysis using data for both H3K27ac and H3K9ac. We found that gene bodies with low H3K27ac EC_{50} scores (more sensitive) are significantly less paused under basal conditions, when compared to the pausing

indices of genes with the most resistance for the association with H3K27ac (least sensitive) (Figure 2.11D, left). In contrast, a similar analysis revealed that the pausing state of RNA Pol II from genes in untreated cells has no statistical significant influence on the dose-dependent changes of H3K9ac (Figure 2.11D, right). Overall, our data show that dose-dependent changes in H3K9ac and H3K27ac by largazole have distinct dose-response behaviors. Genes with low RNA Pol pausing prior to treatment are more sensitive to low dose H3K27 hyperacetylation whereas H3K9 acetylation dose-dependent changes do not seem to be influenced by pausing state.

2.2.8 Largazole induces major changes in the landscapes of histone marks in distal regulatory elements

Increase in RNA Pol II pausing could be the result of a defect in the formation of the preinitiation complex or perturbations of the transcription elongation process (Cheng and Price, 2007; Peterlin and Price, 2006; Yamaguchi et al., 2013). It has been recently shown that enhancer activity has profound impacts on RNA Pol II pausing through production of enhancer RNA (Kim et al., 2015; Li et al., 2013; Schaukowitch et al., 2014). Given our unexpected result of loss of H3K27ac along enhancer-like regions in HCT116 cells treated with largazole (Figure 2.4D), we wondered whether largazole could regulate transcription through the remodeling of enhancer elements. To this end, we measured enhancer associated histone acetylation and RNA Pol II binding as a function of largazole dose in treated cells. In addition, since co-occupancy of H3K4me1 and H3K4me2 are generally associated with functional enhancers (He et al., 2010; Kaikkonen et al., 2013), we performed H3K4me1 and H3K4me2 ChIP-seq of cross-linked nuclear extracts obtained following either vehicle (DMSO), 75 nM, or 300 nM largazole treatment of HCT116 cells.

To help identify active enhancer regions we used published GRO-seq (Allen et al., 2014) and ChIP-seq data for MLL4 and p300 (Hu et al., 2013) in HCT116 cells. We then searched for genomic regions containing overlapping H3K27ac and H3K4me1 peaks (as determined by FStitch and MACS2, respectively) that were not superimposed over annotated transcription start sites. We identified 41,077 inter- and intragenic enhancer locations co-occupied by H3K27ac and H3K4me1 prior to largazole treatment. Gleaning insights from the ChIP-seq data revealed two noteworthy enhancer patterns based on dose-dependent changes induced by largazole. We refer to the first pattern as "enhancer decommissioning", which is characterized by the reduction of H3K27ac, H3K4me1, H3K4me2, as well as RNA Pol II occupancy (Figure 2.12A). In untreated HCT116 cells, most of these enhancer regions display the characteristic bidirectional transcription associated with active enhancers (eRNA). These enhancers are marked by relatively open chromatin as determined by hypersensitivity data, and display high occupancy levels of p300 and MLL4 (not shown). The second pattern is referred to as "enhancer awakening" and is characterized by the dramatic dose-dependent increase in H3K27ac, H3K4me2, and RNA Pol II occupancy as well as the presence of H3K4me1 prior to largazole stimulation (Figure 2.12B). These regions are frequently occupied by MLL4, display unusually high levels of p300, and produce low amounts of eRNA in the basal cellular state (not shown). Detailed analysis of these patterns is described below.

Figure 2.12 Dose-dependent largazole effects on the epigenetic features of distal enhancer elements
Screen shots from Genome Browser (UCSC) showing ChIP-seq and associated signal determined by FStitch (silver) from HCT116 cells targeting H3K9ac (light green) and H3K27ac (orange) starting with untreated cells (DMSO) at the bottom and followed by eight increasing largazole concentrations (4.7 nM to 300 nM). ChIP-seq signal accumulation for total RNA Pol II (dark green), H3K4me2 (pink), and H3K4me1 (yellow) is shown for untreated cells (DMSO) and those treated with either 75 nM or 300 nM largazole concentrations. GRO-seq data from unstimulated HCT116 cells illustrate the presence of nascent transcripts resulting from the plus (red) and negative strand (blue) (Allen et al., 2014). largazole induces both the A) decommissioning and B) activation of transcriptional enhancer regions.

To minimize the number of false-positive deactivated and activated enhancers, we focused on a subset of isolated enhancer regions marked with a single, centered H3K27ac peak in a 20 kb genomic window for further analysis. From the originally identified ~41,000 putative enhancers in untreated cells, we selected 8,667 isolated active enhancers that met the above criteria. Similarly, we selected 3,505 isolated poised enhancers from an initial 10,010 identified elements. To examine epigenetic modifications on enhancer elements as a function of dose, we quantified H3K27ac signal coverage (FStitch) over +/- 1.5 kb enhancer regions centered on overlapping peaks in data from each largazole treatment. *K*-means clustering of the 8,667 isolated active enhancers and subsequent filtering for decreased RNA Pol II accumulation, revealed the presence of 797 largazole-inactivated regulatory elements. We further segregated the deactivated elements into low-dose (416 enhancers) and mid-dose (381 enhancers) affected subsets (Figure 2.13A and 2.13D). The low dose deactivated enhancer cluster displays a high H3K27ac and low H3K9ac signature at the basal state. Low dose largazole treatments erase H3K27ac while the H3K9ac signal retains a low profile. Interestingly, the H3K9ac boundaries associated with these genomic regions undergo a significant expansion with increasing dose stimulations (Figure 2.13A right). Consistent with deactivation of these enhancers, dose-dependent reduction of H3K4me1 and RNA Pol II association were observed and loss of H3K4me2 only occurred at high dose exposure (Figure 2.13B). The high dose deactivated cluster exhibits gradual loss of H3K27ac and a bell-shaped response in H3K9ac changes with increasing largazole exposure (Figure 2.13D). In this cluster of enhancers, H3K4me1 association shows dose-dependent decline while H3K4me2 is unchanged (Figure 2.13E).

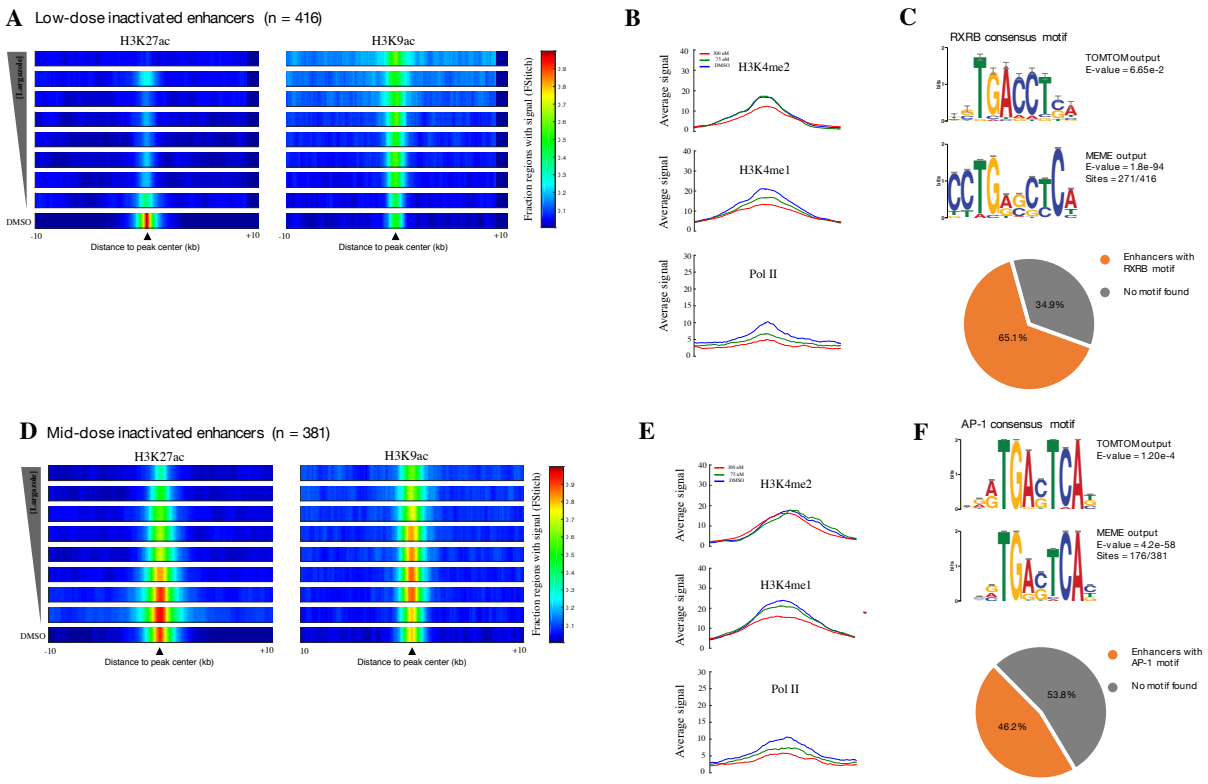


Figure 2.13 Meta-analysis of histone modifications changes, RNA Pol occupancy, and motif enrichment for decommissioned enhancers
A, D) Shown are the fraction of enhancer regions with H3K27ac (left) and H3K9ac (right) signal (FStitch calls) along a +/- 10 Kb distance centered on overlapping peak regions. Peak center locations are indicated by black triangles. Nine ChIP-seq experiments are illustrated with vehicle (DMSO) at the bottom and followed by increasing doses of largazole treatments to a maximum of 300 nM at the top. Fraction of enhancer elements with significant signal (FStitch) for each histone acetylation mark is shown by the heat-color scale: (red) all elements; (green) half of elements; (dark blue) no elements with signal detected. B, E) Average normalized density of ChIP-seq reads for total RNA Pol II, H3K4me1 and H3K4me2 along a +/- 1 Kb distance centered on enhancer regions shown in A. Data from three ChIP-seq experiments are shown; DMSO (blue), 75 nM (green) and 300 nM (red). C, F) Sequence motif associated with the corresponding cluster of poised enhancers. Shown are the determined E-values from the MEME de novo motif finding algorithm and from TOMTOM describing the certainty of the match between the identified motif and the transcription factor database position weight matrices. Pie charts illustrate the percentage of enhancer elements positive for the identified consensus motif.

A similar analysis on the selected awakened poised enhancers yielded low-dose (688 elements) and high-dose (914 elements) stimulated subsets (Figure 2.14A and 2.14D). The high dose cluster exhibited a largazole dose-dependent enrichment of H3K27ac, RNA Pol II association and to a lesser degree it also accrued H3K9ac signal. There is only a slight increase in H3K4me1 in this group. In contrast, H3K4me2 signal was barely detectable with DMSO and

75 nM largazole treatment but elevated drastically upon treatment with 300 nM largazole which correlates with the dose-dependent rise of H3K27ac (Figure 2.14B). The low dose induced cluster showed a gradual increase in H3K27ac and RNA Pol II but a fluctuating H3K9ac level as largazole dose was increased (Figure 2.14D & 2.14E). H3K4me2 was unchanged and H3K4me1 displayed a slight decrease at 300 nM largazole.

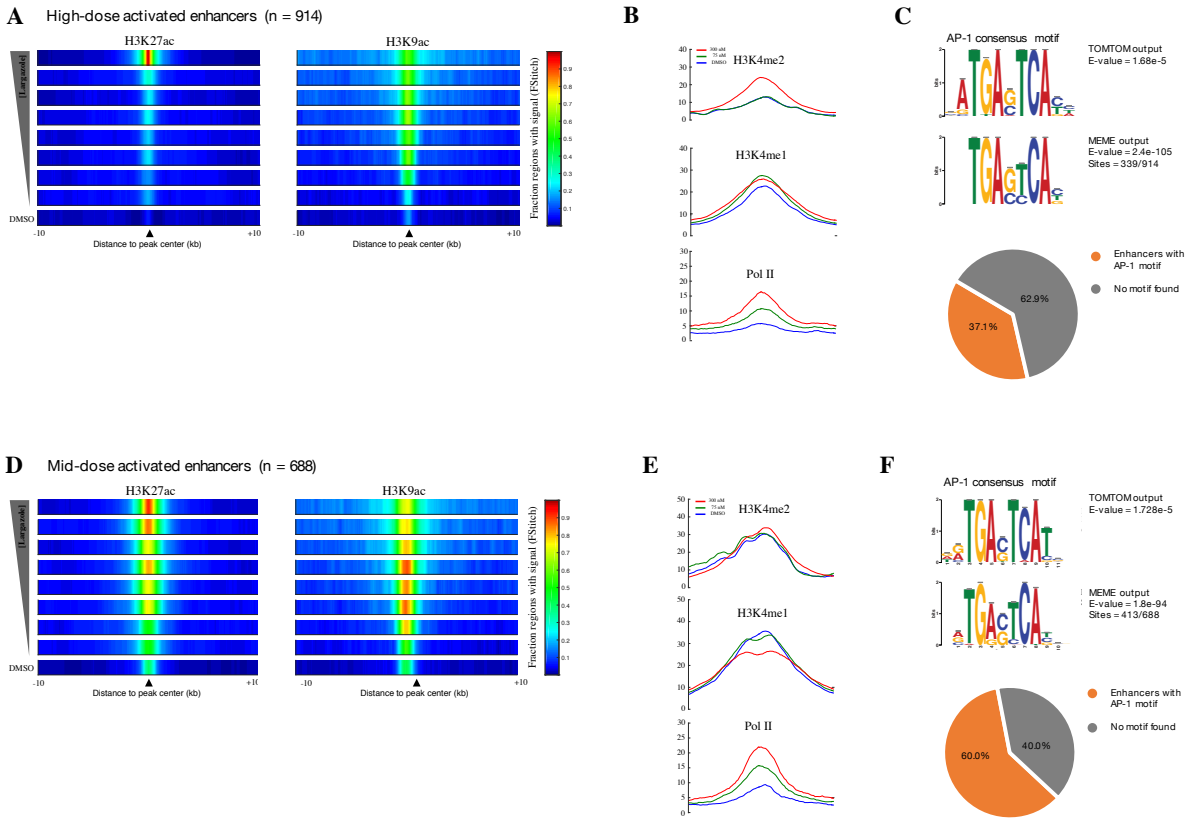


Figure 2.14 Meta-analysis of histone modifications changes, RNA Pol occupancy, and motif enrichment for activated enhancers
A, D) Shown are the fraction of enhancer regions with H3K27ac (left) and H3K9ac (right) signal (FStitch calls) along a +/- 10 Kb distance centered on overlapping peak regions. Peak center locations are indicated by black triangles. Nine ChIP-seq experiments are illustrated with vehicle (DMSO) at the bottom and followed by increasing doses of largazole treatments to a maximum of 300 nM at the top. Fraction of enhancer elements with significant signal (FStitch) for each histone acetylation mark is shown by the heat-color scale: (red) all elements; (green) half of elements; (dark blue) no elements with signal detected. B, E) Average normalized density of ChIP-seq reads for total RNA Pol II, H3K4me1 and H3K4me2 along a +/- 1 Kb distance centered on enhancer regions shown in A. Data from three ChIP-seq experiments are shown; DMSO (blue), 75 nM (green) and 300 nM (red). C, F) Sequence motif associated with the corresponding cluster of poised enhancers. Shown are the determined E-values from the MEME de novo motif finding algorithm and from TOMTOM describing the certainty of the match between the identified motif and the transcription factor database position weight matrices. Pie charts illustrate the percentage of enhancer elements positive for the identified consensus motif.

A striking feature associated with the awakened poised enhancers is that under basal conditions they tend to have high p300 occupancy yet minimal or absent H3K27ac signal (Figure 2.15). The presence of H3K4me1 and absence of H3K27ac in untreated HCT116 cells suggest that this subset of enhancers is under a poised state but likely primed for prompt activation. Our results suggest that HDACs are probably actively involved in maintaining the poised state and largazole inhibition of histone deacetylases tip the balance in favor of H3K27 acetylation. Taken together, this shows that largazole acts through inhibition of HDAC targets both in the deactivation and re-activation of distinct classes of enhancer elements that can be discernible by their dose sensitivity.

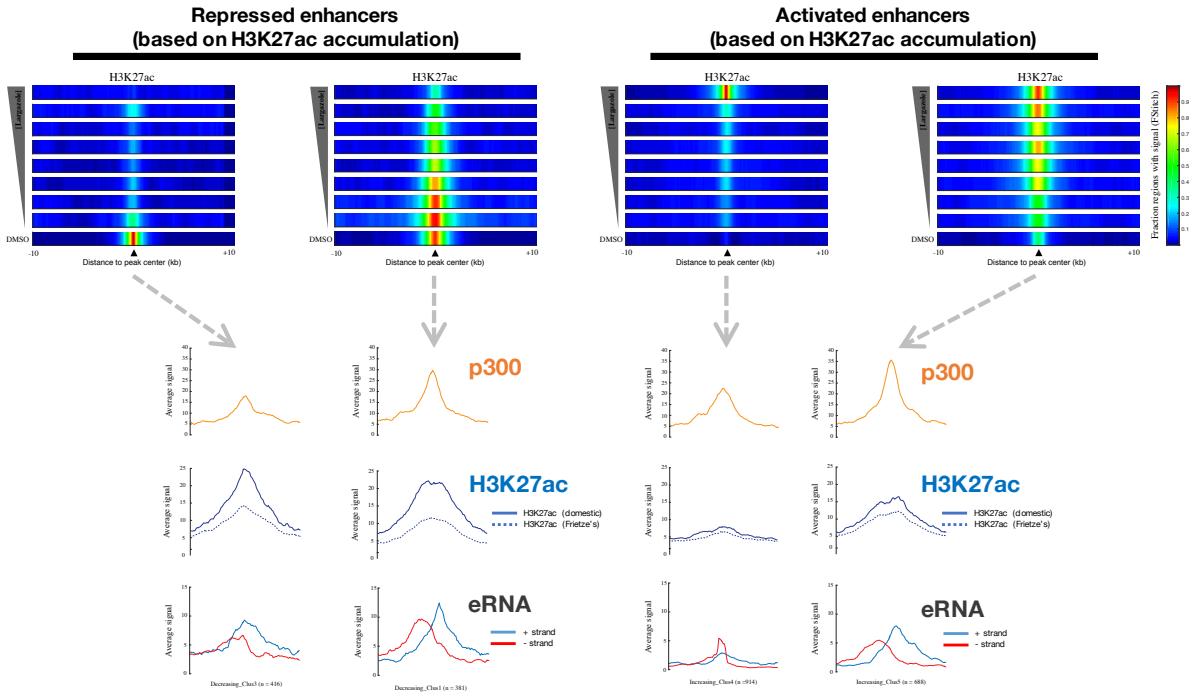


Figure 2.15 p300 occupancy levels at active and poised enhancers under basal cellular conditions
Four clusters of enhancers identified based on their largazole-induced functional state and dose-response: largazole-repressed enhancers at low-dose (n = 416), mid-dose (n = 381) and largazole-activated enhancers at high-dose (n = 914) and mid-dose (n = 688). Top panel, shown are the fraction of enhancer regions with H3K27ac signal (FStitch calls) along a +/- 10 Kb distance centered on overlapping peak regions. Peak center locations are indicated by black triangles. Nine ChIP-seq experiments are illustrated with vehicle (DMSO) at the bottom and followed by increasing doses of largazole treatments to a maximum of 300 nM at the top. Bottom panel, average normalized density of ChIP-seq reads from unstimulated HCT116 cells for p300 (orange) (Hu et al., 2013), H3K27ac (solid blue from domestic and dotted blue from (Frietze et al., 2012)), and nascent RNA (blue and red) (Allen et al., 2014) along a +/- 1 Kb distance centered on enhancer regions shown on top.

2.2.9 Epigenetically remodeled enhancer regions are heavily enriched with the recognition motif for signal induced transcription factors

HDAC inhibitors have been implicated in the differentiation of transformed cells and in the redirection of embryonic cell lineages (Haumaitre et al., 2008; Marks et al., 2000; Svechnikova et al., 2008). Recent studies suggest that a relatively small combination of lineage-determining factors act in concert to prime a large number of enhancer-like regions during the differentiation of a particular cell type (Heinz et al., 2010). Moreover, during embryonic stem cell differentiation the deactivation of pluripotency related genes is mediated by demethylation of H3K4me1 and deacetylation of H3K27ac at the associated enhancers (Whyte et al., 2012). To gain a better understanding of the mechanism of selection for enhancer deactivation and re-activation under different largazole dose stimulations, we searched for enriched DNA recognition motifs within the enhancers in each cluster. We found that re-activated enhancers were highly enriched for the canonical motif recognized by the transcription factor AP-1 (activator protein-1). Specifically, 60.0% of the low-dose and 37.1% of the high-dose newly deployed enhancers contained the AP-1 recognition motif (Figure 2.14C & 2.14F). Similarly, the mid-dose decommissioned enhancers were also enriched with the same motif (46.2%). The cluster with low-dose deactivated enhancers showed a high percentage of enrichment (65.1%) for a motif similar to the DNA response element of the retinoic acid receptor RXRB (Figure 2.13C). The AP-1 complex can induce or repress gene transcription and it is involved in a wide range of cellular process, including cell proliferation, death, survival, and differentiation. Its diverse functionality stems from an inherited structural complexity. AP-1 can be found as a hetero- or homo-dimer complex composed of members of the Jun, Fos, Maf, and ATF sub-families. More often, c-Jun is a positive regulator of cell proliferation, whereas JunB has the

opposite effect (Shaulian and Karin, 2001, 2002). In agreement with this model, we found the accumulation of most of these transcripts changing systematically in response to largazole dose treatments. Particularly, c-Jun mRNA accumulation increased dose dependently starting in cells treated with an 18.75 nM largazole while the JunB transcript displayed a significant accumulation only in cells treated with largazole concentration of 37.5 nM or higher (Figure 2.16). Taken together, our findings suggest that restructured enhancers might be primarily under the control of signal induced transcription factors such as AP-1 complex which are likely responsible for transcriptional reprogramming of HCT116 cells stimulated with largazole.

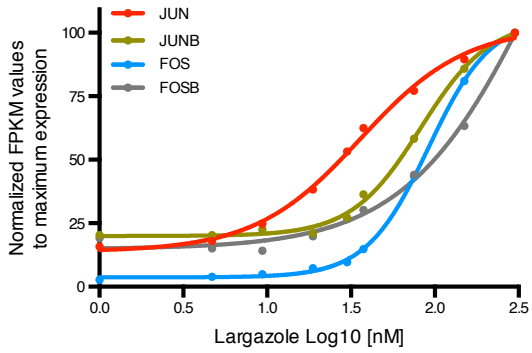


Figure 2.16 Largazole stimulates transcriptional activation of genes coding for the protein members of the AP-1 complex
mRNA accumulation levels of members of the AP-1 transcription factor complex from HCT116 cells treated for 16 hours with the indicated largazole concentration. Transcript levels are shown as fragments per kilobase of exon per million mapped reads (FPKM) normalized to maximum expression value.

2.2.10 Gene ontology analysis of differentially expressed genes might explain the cancer-specific lethality of largazole

In order to gain a better understanding of the biological implications of differentially expressed transcripts as a function of dose, we performed DESeq analysis on mRNA-seq data for each pairwise comparison of treatments (i.e. DMSO vs 4.68 nM largazole treatment). We identified 1,245 unique differentially expressed transcripts. For the most part, transcripts that are

differentially expressed in cells treated with low doses of largazole maintain both incremental changes and consistent directionality in cells stimulated with higher doses (Figure 2.17A). Dose dependent analysis of differentially expressed transcripts revealed a general mRNA accumulation trend where cells exposed to low nanomolar concentrations of largazole resulted mostly in the up-regulation of gene transcripts whereas mid to high doses primarily triggered a decreased in mRNA accumulation (Figure 2.17B).

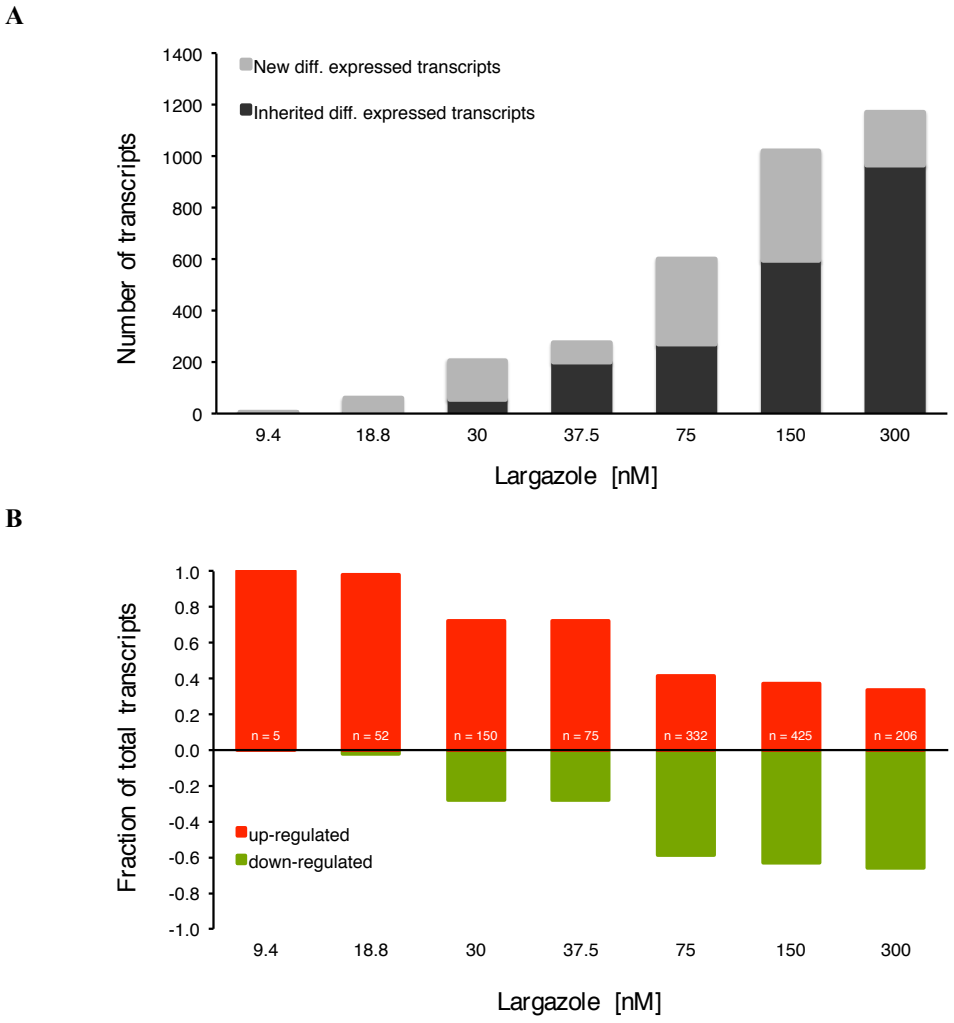


Figure 2.17 Low doses of largazole treatment in HCT116 cells induce mRNA upregulation while high doses lead primarily to transcript levels attenuation
A) Total number of genes differentially expressed at each largazole dose treatment showing newly differentially expressed in grey and transcripts that inherited from a lower dose in black. B) Differentially expressed transcripts unique to each largazole dose treatment based on DESeq analysis with an adjusted p-value cutoff of less than 0.1. Transcripts are shown as a fraction of total elements per dose where those that are upregulated (red) are plotted above the zero line and below those that are downregulated (green).

We performed Gene Ontology (GO) enrichment analysis using the online based software DAVID (Huang et al., 2009b, 2009c). We found agreement between the enriched GO terms based on differentially expressed transcripts and previously reported biological activities of largazole. For example, largazole is known to repress the expression of vascular endothelial growth factor genes (VEGFs) and to efficiently inhibit angiogenesis (Liu et al., 2013b). As shown in table 2.2, blood vessel development is one of the top categories enriched for differentially expressed transcripts in cells treated with 30 nM, 37.5 nM, and 75 nM largazole concentrations. Overall, we found two "cell death" terms (GO:0010941 and GO:0008219) consistently throughout the six largazole treatments analyzed and we noticed a shift from either nucleosome assembly (GO:0031497 and GO:0006334) or regulation of transcription (GO:0006357) terms in low doses to cell cycle (GO:0007049) and mitosis (GO:0007067) terms in mid- to high- largazole cell treatments. This suggests that several of these genes can be related to largazole's ability to induce cell cycle arrest and apoptosis and are likely influenced directly or indirectly by components of the AP-1 complex.

Table 2.2 Gene Ontology categories based on the identity of differentially expressed mRNAs for each largazole drug treatment

Gene Ontology (GO) analysis for differentially expressed transcripts for six largazole drug treatments (18.75 nM to 300 nM) using the database for annotation, visualization, and integrated discovery (DAVID) bioinformatics tools (Huang et al., 2009b, 2009c). Shown are selected top GO biological processes based on their enrichment p-value and associated number of genes.

GO term [18.75 nM]	p-value
Response to stimulus (n=5)	6.54E-03
Regulation of hormone levels (n=4)	1.21E-02
Positive regulation of cell adhesion (n=3)	1.55E-02
Cell motion (n=6)	1.68E-02
Homotypic cell-cell adhesion (n=2)	2.20E-02
Regulation of FGF signaling (n=2)	2.20E-02
Cell death (n=7)	2.51E-02

GO term [30 nM]	p-value
Nucleosome assembly (n=7)	3.24E-04
Blood vessel development (n=11)	3.82E-04
Regulation of cell migration (n=7)	1.10E-02
Negative regulation of signal transduction (n=8)	1.12E-02
Cell death (n=16)	1.29E-02

GO term [37.5 nM]	p-value
Blood vessel development (n=15)	1.49E-05
Nucleosome assembly (n=8)	2.26E-04
Positive regulation of TGFb signaling (n=4)	7.93E-04
Response to organic substance (n=23)	8.80E-04
Response to wounding (n=18)	2.04E-03
Cell death (n=20)	9.22E-03

GO term [75 nM]	p-value
Response to organic substance (n=49)	1.05E-06
Regulation of transcription (n=48)	3.06E-06
Positive regulation of metabolic process (n=53)	5.64E-06
Blood vessel development (n=22)	3.96E-05
Regulation of cell proliferation (n=47)	5.12E-05
Cell death (n=40)	8.40E-04

GO term [150 nM]	p-value
Cell cycle (n=90)	2.46E-12
Positive regulation of metabolic process (n=83)	1.07E-07
M phase (n=42)	3.01E-07
Response to organic substance (n=71)	5.80E-07
Positive regulation of transcription (n=43)	2.82E-06
Mitosis (n=30)	5.43E-06
Regulation of cell death (n=73)	1.21E-05

GO term [300 nM]	p-value
Cell cycle (n=93)	3.17E-10
Response to organic substance (n=85)	4.61E-09
Apoptotic mitochondrial changes (n=12)	1.09E-06
M phase (n=44)	1.63E-06
Nuclear division (n=33)	3.75E-06
Mitosis (n=33)	3.75E-06
Apoptosis (n=66)	4.07E-06
*Regulation of cell death (n=82)	6.75E-06

2.3 Discussion

As a potent isoform-selective HDACI, largazole induces dose-dependent inhibition of cell proliferation and transcriptional reprogramming. The effects of HDACs on transcriptional regulation have been long recognized but the underlying mechanisms are quite complex. Multiple isoforms of HDACs have overlapping functions and associate with distinct molecular complexes that function in a genomic and cellular context dependent manner. This poses significant challenges to the elucidation of the mechanistic link between HDACs and their biological activities. In eukaryotic cells, gene expression can be regulated through either accessibility of proximal promoters to the transcriptional machinery or by modulating pausing of RNA Pol II during elongation (Gilmour and Fan, 2009). In *Drosophila* cells, HDAC inhibition by both TSA and SAHA contribute to histone H3 acetylation at promoters and downstream regions. This event stimulates both transcription initiation and elongation (Zhao et al., 2005). Conversely, GRO-seq analysis of human BT474 cells treated with the pan-HDACIs, showed that TSA or SAHA induced a decrease of transcription along gene bodies without affecting nascent transcript production at the corresponding promoters (Kim et al., 2013). A more recent study suggests that TSA or SAHA perturbs transcription elongation by preventing eviction of NELF at promoters and loss of eRNA synthesis at some enhancers (Greer et al., 2015). In this study we perform RNA-seq analysis with increasing doses of largazole. Our results reveal a clear trend in terms of transcriptional activation and repression for largazole. Low doses of largazole (9.4 nM and 18.8 nM) mostly induce transcriptional activation but as the dose of largazole increases the fraction of up-regulated transcripts decreases (Figure 2.17B). At 300 nM, more genes are repressed than activated. Therefore, transcriptional reprogramming by HDAC inhibition is dose-dependent.

Dose-dependent transcriptome changes correlate with the biological responses of cancer cells. It has been shown previously that a low dose of largazole induces cell growth arrest at the G1 phase of the cell cycle while a high dose of largazole causes G2/M arrest and apoptosis (Liu et al., 2010). Based on our results, it is tempting to speculate that transcriptional activation at low largazole doses may contribute to cell cycle arrest at G1 and that the profound transcriptional repression observed upon treatment with a high largazole dose is linked to G2 arrest and apoptosis. Since most therapeutic drugs are administrated just below the maximum dose tolerance (MTD), HDACI-induced transcriptional repression is probably highly relevant to their therapeutic benefits and also undesirable toxicity. Recent clinical success in treatment of ER positive breast cancer with Cdk4/6 inhibitor Palbociclib (Pfizer) sparks renewed interest in developing inhibitors that block G1 to S transition and promote cell differentiation (Turner et al., 2015). One implication of our study is that low doses of HDACI could also be an effective yet unexplored treatment strategy, especially in developing combination therapies.

To understand the fundamental mechanisms underlying HDAC inhibitor induced transcription activation and repression, we investigated dose-dependent changes in acetylation patterns of two lysine residues in histone H3, namely H3K9 and H3K27. In response to largazole treatment both H3K9ac and H3K27ac show dose-dependent elevation and spread in the gene body of actively transcribed genes. Despite these spectacular changes in the gene body acetylation, it hardly correlates with up- or down-regulation of transcription. Our observations are in agreement with previous studies with other HDACIs (Greer et al., 2015; Rafehi et al., 2014; Wang et al., 2009b). This prompted us to search for changes in the distal regulatory elements that may control transcriptional activity and led us to the findings that a subset of enhancer elements is awakened from the latent state to acquire hallmarks of active enhancers.

Furthermore, our studies also revealed a subset of enhancer elements that are decommissioned under largazole treatments.

Accumulating evidence now supports the notion that there are active, poised and latent enhancers in the genome defined by their distinct histone marks (Heinz et al., 2015; Shlyueva et al., 2014). Poised enhancers bear the features of H3K4me1 or H3K4me2 histone marks, absent or low acetylation of H3K27 and minimal association with RNA Pol II. It is still unclear how poised enhancers are maintained and reactivated. Based on our results we propose a model in which both HATs and HDACs occupy the poised enhancer region but classical HDACs maintain a repressed state. Once inhibited by largazole, the balance shifts toward acetylation of H3K9/27, critical for subsequent enhancer activation. Several lines of evidence support this hypothesis. In macrophages, the NCoR1/HDAC3 corepressor complex is recruited to promoter regions bearing AP-1 binding sites and acts as a transcription factor checkpoint likely mediating deacetylation of histone tails required for transcription activity (Ogawa et al., 2004). These observations imply that HDAC complexes could very well be recruited to particular enhancer elements for active histone deacetylation and repression of specific gene transcripts. Indeed, in $CD4^+$ T cells, class I HDACs (HDAC2, and 3) and several HATs (p300, CBP, PCAF, MOF, and Tip60) can be bound to the same intergenic regions at high frequency, suggesting a dynamic histone acetylation remodeling at these locations (Wang et al., 2009b). Interestingly, under basal cellular conditions we found high levels of p300 ChIP-seq signal at intergenic regions that will become awakened enhancers under largazole stimulation. However, these genomic locations do not have detectable H3K27ac signal (Figure 2.15). Given largazole has superb potency against class I HDAC enzymes, we speculate that this class of HDACs are involved in maintaining poised enhancers in the repressive state.

Loss of H3K27 acetylation at enhancer elements upon largazole treatment is rather counterintuitive. It happens more often with higher doses of largazole and correlates with gene repression. Consistent with previous studies with other HDACIs, RNA Pol II shows a dose dependent increase in pausing which may be a result of defects in the transition from initiation to elongation or blockage in elongation (Greer et al., 2015; Wang et al., 2009b). The observed decreased of H3K27ac at enhancer regions could be an indirect effect of HDAC inhibition treatment. One possible explanation is that at high doses, largazole enables suppression of HATs through the stimulation of the Polycomb complex (PC). It has been shown that PC is associated with many active promoters that exhibit Pol II pausing (Enderle et al., 2011; Schaaf et al., 2013). In *Drosophila*, PC and CREB-binding protein (CBP) co-occupy many genomic sites including enhancers and the promoter regions of active and repressed genes (Goodman and Smolik, 2000; Tie et al., 2014, 2015). Recruitment of PC to actively transcribed genes is independent of H3K27me3. It has been proposed that PC binds directly to CBP, HAT, and AIL and inhibits histone acetylation activity thereby lowering H3K27ac levels (Tie et al., 2015).

In summary, our genome-wide dose-response analysis of transcriptome and histone signatures revealed new target specificity of largazole in transcriptional reprogramming. Our studies provide a more mechanistic explanation of the effecto of HCDAIs on gene expression. Future studies focusing on dynamic changes of histone signatures and more comprehensive profiling of histone marks should unravel more insights into remodeling the enhancer landscapes that link to therapeutic responses *in vivo* and ultimately uncover predictive biomarkers.

2.4 Materials and methods

2.4.1 Cell culture and largazole treatment

HCT116 colorectal carcinoma cells were kindly provided by Dr. Joaquin Espinosa, University of Colorado at Boulder. Cells were cultured in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (Sigma), 1% penicillin streptomycin, and 1 % GlutaMAX (Invitrogen) and maintained at 37°C and 5% CO_2. Prior to treatment, HCT116 cells were grown in complete media and passaged for 3 consecutive days. Cells were treated with the indicated largazole concentration or equivalent amount of vehicle (DMSO) at 70% confluency and harvested after 16 hours for all ChIP-seq experiments as well as for immunoblotting assays.

2.4.2 RNA extraction and library preparation

Total RNA was extracted from 16 hours treated HCT116 cells using TRIzol reagent (Life Technologies) according to the manufacturer's protocol. The concentration of each sample was quantified using the Qubit™ 3.0 Fluorometer (Thermo Fisher), and integrity was measured on an Agilent Bioanalyzer 2100 (Agilent Technologies). The Illumina TruSeq RNA Sample Prepartaion kit (Illumina) was used to generate the RNA sequencing libraries. Briefly, mRNA was purified from 2.5 ug total RNA from each sample, fragmented, and converted to double-stranded cDNA with the use of modified oligo(dT) primers. Sequencing barcodes were ligated to the cDNA fragments, and the resulting fragments were amplified using PCR. The final lengths of oligos from each library were validated on an Agilent Bioanalyzer 2100.

2.4.3 Sequencing

Libraries were quantified using the Qubit™ 3.0 Fluorometer and sequenced at the Next-Generation Sequencing Facility at the University of Colorado BioFrontiers Institute. All

sequencing libraries were multiplexed and sequenced on an Illumina HiSeq 2000 sequencing system (Illumina). For each sample we obtained between 14.72 and 19.52 million 50-bp reads.

2.4.4 RNA mapping and normalization

Reads were trimmed to a final length 43bp and mapped to human genome 18 (RefSeq) using Bowtie and TopHat (Trapnell et al., 2009), a short read aligner that is capable of predicting exon-exon splice junctions. After mapping, alignment files were processed using SAMtools (Li et al., 2009). Using Cuffdiff we counted the total number of sequencing reads that aligned to each putative gene model in the human genome. To determine which genes were differentially expressed, we used the R package DESEq (Simon and Wolfgang, 2010).

2.4.5 Immunoblotting, antibodies, and signal quantification

Western blots were carried out using standard protocols. Briefly, HCT116 cells were grown, treated, and harvested as previously mentioned. Nuclear protein lysates were separated by SDS-PAGE and transferred to GVS nitrocellulose 0.22 micron membranes. Blots were probed with primary antibodies, followed by peroxidase-conjugated secondary antibodies (GE Healthcare Life). Signal for all immunoblots was acquired using the ImageQuant LAS 4000 biomolecular imager (GE Healthcare LS) with an average exposure of 30 seconds. For the quantification of the bands, a custom fitting algorithm was scripted in MATLAB (Matworks R2015a) using the unmodified .tiff files. EC_{50} plots were performed using Prism 6 (GraphPad Software). The equation for manual fitting of EC_{50} plots is $y = max + ((min-max)/(1 + (X/EC_{50})^{n}))$, where X is the concentration of inhibitor, and n is the Hill coefficient. Antibodies used are as follows: H3K9ac (abcam, cat. # ab4729); H3K27ac (abcam, cat. # ab4729); H3K4me1 (abcam, cat. # ab8895); H3K4me2 (abcam, cat. # ab7766); H3K4me3 (abcam, cat. # ab8580); total H3 (abcam, cat #1791).

2.4.6 Gene Ontology (GO) and network visualization analyses

GO analyses were completed using DAVID bioinformatics tools (Huang et al., 2009b). We used gene identities from accumulated differentially expressed transcripts at each largazole concentration. Resulting GO categories were sorted based on increasing *p* values. Because GO categories can often be redundant, we selected a single category as representation and reported the most significant terms (Table 2.2).

2.4.7 Chromatin immunoprecipitation

HCT116 cells were treated with largazole or vehicle for 16 hours and cross-linked with 1% formaldehyde for 15 minutes at room temperature (25°C). Cells were washed two times with PBS and membranes ruptured in hypotonic buffer (50 mM NaCl, 1% NP-40 alternative, 2 mM EDTA, 10 mM Tris, 1 mM DTT, 2 mM EDTA, 1X protease inhibitor cocktail (Roche # 04693124001). The cell nuclei were recovered by centrifugation and resuspended in lysate buffer (150 mM NaCl, 0.5% Triton X-100, 2 mM EDTA, 0.1% SDS, 20 mM Tris, 1 mM DTT, 1X protease inhibitor cocktail). Resuspended samples were sonicated for 25 cycles (30s 'on' at high level and 30s 'off' per cycle) using a Bioruptor (Diagenode; Denville, NJ, USA) and spun for 10 minutes at 16,000x g in a microcentrifuge. Samples were incubated for 5 hours at 4°C with 5 to 20 μg of antibodies and 20 μl of 50% slurry with protein A beads (Millipore; Billerica, MA, USA). The immunoprecipitated chromatin was then recovered and DNA purified using phenol chlorophorm extraction. Sequencing libraries were prepared using an Illumina ChIP-Seq DNA Sample Prep Kit (cat. # IP-102-1001), with a starting sample varying from 2 to 20 ng of DNA isolated from the immunoprecipitation step. Antibodies used are as follows: RNA Pol II

(Santa Cruz sc-899 lot # K0111); H3K9ac (abcam, cat. # ab4729); H3K27ac (abcam, cat. # ab4729); H3K4me1 (abcam, cat. # ab8895); H3K4me2 (abcam, cat. # ab7766).

2.4.8 ChIP-seq mapping and normalization

ChIP-seq datasets were aligned using Bowtie mapping software version 0.12.7 (Langmead et al., 2009). To maintain the same read length across all experiments, H3K4me1 and H3K4me2 ChIP-seq raw datasets (fastq files) were trimmed to 50bp using FASTX-toolkit (version 0.0.13.2) (http://hannonlab.cshl.edu/fastx_toolkit/). All reads were mapped to the hg18 reference human genome with a number of base pairs mismatch not greater than 2 (96% sequence match). We used SAMtools version 0.1.19 (Li et al., 2009) to generate a sorted pileup format of the aligned reads. Reads were then extended from the 3'–end to a final length of 150bp. For each experiment, genome coverage bed graph files were generated using BEDTools2 version 2.25.0 (Quinlan and Hall, 2010) and then normalized by multiplying the read density times 100 and dividing by the total number of mapped reads. Normalized bed graph files were subsequently converted to bigwig files and uploaded to UCSC Genome Browser for visualization.

We downloaded ChIP-seq data for p300, MLL4, as well as input from HCT116 cells previously published (Hu et al., 2013), from the GEO database accession number GSE1176. In addition, we also acquired published GRO-seq data for HCT116 cells from the GEO database accession number GSE53964 (Allen et al., 2014). Raw ChIP-seq data for p300, MLL4, as well as the GRO-seq data were processed in the same manner as mentioned above.

2.4.9 Identification of ChIP-seq signal

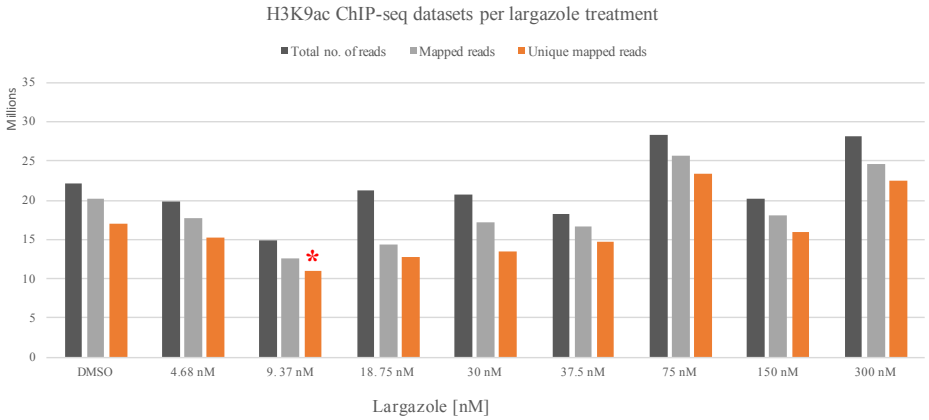
2.4.9.1 H3K4me1, H3K4me2, RNA Pol II, p300, and MLL4

With the exception of H3K9ac and –K27ac, signal analyses for all ChIP-seq datasets experiments were performed using MACS2 version 2.1.0.20150731 (Zhang et al., 2008) under default settings and a *p*-value cutoff of 1e-05. We used --broad -g hs --keep-dup=auto -p 1e-5 -m 10 200 --bw 200 and selected broad peak calls.

2.4.9.2 H3K9ac and H3K27ac

FStitch algorithm was used to identify genomic regions enriched with H3K9ac and –K27ac signal from ChIP-seq experiments. In order to acquire uniform FStitch signal calls across experiments targeting the same acetylated lysine, we determined the minimal number of unique reads found in datasets for H3K9ac as well as in those for H3K27ac (Figure 2.18A & 2.18B respectively). Based on these numbers, we randomly subsample 12,844,004 unique reads from all H3K9ac ChIP-seq experiments and 9,122,018 unique reads from all nine H3K27ac ChIP-seq datasets. For H3K9ac ChIP-seq data analyses, we used 20 genomic regions from untreated HCT116 H3K9ac ChIP-seq data as FStitch-training genomic locations (Table 2.3). In a similar manner, we used 19 genomic regions from H3K27ac under basal experimental conditions as FStich training parameters (Table 2.4). Segmentation analysis for all ChIP-seq experiments targeting the same lysine on histone H3 were conducted using the output parameters gathered from the training sessions. The same analysis was performed on the input experiment and any resulting signal was subtracted from all ChIP-seqs.

A



B

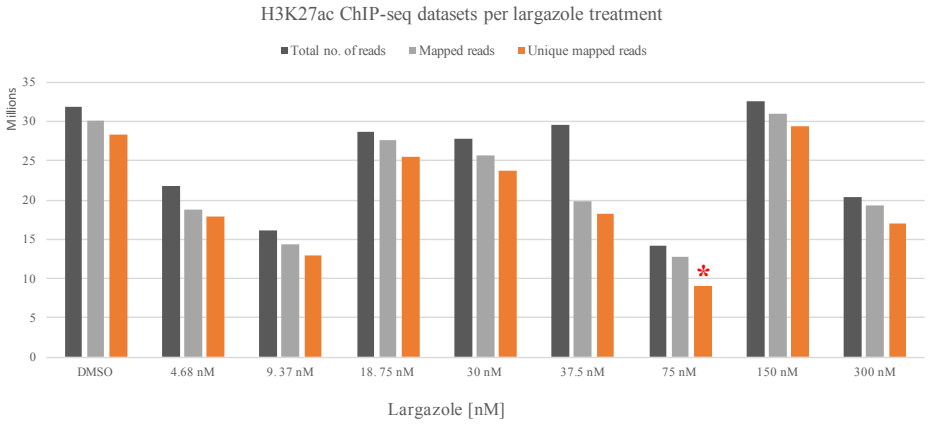


Figure 2.18 Number of mapped reads from individual ChIP-seq experiment targeting H3K9ac and H3K27ac
Shown are total number of reads (grey), number of mapped reads to hg18 (silver) based on criteria mentioned in materials and methods, number of unique mapped reads (orange). Red asterisk denotes experiment with the minimum number of unique mapped reads.

Table 2.3 H3K9ac training genomic coordinates for FStitch analysis
Table illustrates genomic coordinates (hg18) utilized for FStitch training using H3K9ac ChIP-seq from unstimulated HCT116 cells, assuming enrichment (signal state = 1) or background levels (signal state = 0).

chromosome number	start	end	signal state
chr1	1	609	1
chr1	610	17779	0
chr1	17780	20366	1
chr1	20367	27926	0
chr1	27927	28863	0
chr1	28864	358752	0
chr1	359685	431168	0
chr1	431169	432700	1
chr1	432701	448575	0
chr1	448576	450082	0
chr1	450083	513397	0
chr1	513398	514090	0
chr1	514091	529796	0
chr1	529797	531292	1
chr1	531293	554227	0
chr1	554228	560352	1
chr1	560353	662915	0
chr1	664185	702501	0
chr1	848578	852147	1
chr1	887232	890566	0

Table 2.4 H3K27ac training genomic coordinates for FStitch analysis
Table illustrates genomic coordinates (hg18) utilized for FStitch training using H3K27ac ChIP-seq from unstimulated HCT116 cells, assuming enrichment (signal state = 1) or background levels (signal state = 0).

chromosome number	start	end	signal state
chr1	1	526	1
chr1	527	17750	0
chr1	17751	21131	1
chr1	21132	27357	0
chr1	29318	79623	0
chr1	80395	431321	0
chr1	431322	432526	1
chr1	432527	530041	0
chr1	530042	531681	1
chr1	531682	554227	0
chr1	554228	560240	1
chr1	560241	702800	0
chr1	702801	705745	1
chr1	705746	750847	0
chr1	753988	829258	0
chr1	829259	832773	1
chr1	832774	845885	0
chr1	845886	852054	1
chr1	863540	868379	1

2.4.10 Defining proximal promoters and positive association with RNA Pol II

To select for genes bound by RNA Pol II at transcription start sites, we defined proximal gene promoter regions as 100bp +/- from annotated TSSs using the January 2016 UCSC RefSeq gene assembly (hg18). RNA Pol II ChIP-seq signal was determined using MACS2 (version

2.1.0.20150731) narrow peak analysis based on default settings and a *p* value cutoff equal to 1e-05 (Zhang et al., 2008). Using merged peak-signals of fragments within a 1kb range resulting from MACS2 analysis, we identified 24,619 proximal gene promoter regions positively associated with RNA Pol II in untreated HCT116 cells. Because many annotated genes contain multiple isoforms associated with a single TSS, we selected for the longest annotated gene versions and for genes which bodies did not overlap with other genes (n = 8765). From this list, we excluded genes which associated TSSs were within 2kb from neighboring genes (n = 2130), genes which annotated lengths are smaller than 3kb (n = 303), genes that contained intragenic enhancer elements (based on H3K27ac and H3K4me1 co-occupancy) (n = 3923), as well as genes that displayed multiple internal TSSs occupied by RNA Pol II (n = 58). Using this method, we identified 2,352 protein coding genes bound by RNA Pol II at the associated TSS and suitable for pausing index assessment.

2.4.11 Pausing index calculation

Calculations were performed as in (Rahl et al., 2010). For the selected genes (n = 2352), we defined promoter regions from –30 to +300 relative to the TSS and the gene body extending from +300bp to the end of the gene annotation (Figure 2.8). RNA Pol II accumulation at promoters and gene bodies was determined using unique mapped reads from RNA Pol II ChIP-seq experiments of untreated (15,129,717), and largazole treated HCT116 cells with 75 nM (15,993,361), and 300 nM (13,869,502) concentrations. Read density for promoter and gene body windows were calculated by dividing the number of unique reads by the total base pairs associated with each specified window. Pausing index was assigned to each gene from the ratio between RNA Pol II density in the promoter region to that of the gene body.

2.4.12 De novo motif analysis

For de novo motif discovery, we used MEME (Bailey et al., 2009). Analysis were performed with a search window of 800 and 500bp flanking the center enhancer elements associated with the defined cluster. The reported E-value is the output of the MEME de novo motif finding algorithm. To identify related transcription factors, each identified motif was input to TOMTOM version 4.11.1 (Gupta et al., 2007) using motif database JASPAR DNA CORE (2016) or HUMAN DNA HOCOMOCO (v10). We also report the E-value describing the certainty of the match between the identified de novo motif and the database position weight matrices. The images were prepared using Adobe Illustrator CS6 or Photoshop CS6.

2.4.13 Identification of transcriptional enhancer elements

We first determined H3K27ac ChIP-seq signal (FStitch), as mentioned above, resulting from unstimulated and largazole treated cells. We performed fragment intersect analyses to extract genomic regions with overlapping H3K27ac and H3K4me1 accumulation. To further define the boundaries of enhancer regions, we trimmed the co-occupied regions using MACS2 broad peak calls gathered from H3K4me2, RNA Pol II, and MACS2 narrow peaks from p300 (Hu et al., 2013) ChIP-seq data gathered from unstimulated HCT116 cells. We then eliminated all genomic regions which coordinates overlapped with annotated transcription start sites based on the January 2016 UCSC RefSeq gene assembly (HG18). This led to identification of 41,077 putative enhancer elements in unstimulated HCT116 cells.

2.4.14 *K*–means clustering of H3K27ac signal along enhancer regions

K–means clustering was performed on the enhancer data set referred as "isolated enhancers" (n = 12,172) characterized by a single H3K27ac peak, co-occupied by H3K4me1, centered along a 20kb genomic region, either under basal cellular conditions (n = 8,667) or

resulting from stimulation with 300 nM largazole treatment (n = 3,505). To this end, we used H3K27ac FStitch calls from all nine ChIP-seq experiments along +/- 1kb distance centered on overlapping peak regions (H3K27ac, H3K4me1 (MACS2 BP), and H3K4me2 (MACS2 BP), RNA Pol II (MACS2 BP) or p300 (MACS2 NP) when present). Two filters were applied on the analyzed enhancer list. First, the *K*–means clusters were selected based on two general H3K27ac signal trends, decreasing or increasing under largazole treatments. Second, an additional filtered was applied on these clusters based on RNA Pol II normalized read density patterns; selecting enhancer regions with both decreasing H3K27ac ChIP-seq signal in concomitance with an overall decreasing in RNA Pol II occupancy (based on DMSO, 75 nM, and 300 nM data) or enhancer elements with increasing H3K27ac ChIP-seq signal accompanied by the systematic increase of RNA Pol II binding.

2.4.15 STAT1 knockdown using shRNA

The STAT1 lentiviral shRNA vector (TRCN0000280021) was obtained from the Functional Genomics Facility at the University of Colorado at Boulder. To generate lentiviral particles, $5x10^6$ HEK293T cells were seeded per 10 cm plates, grown overnight, and followed by cotransfection of lentiviral shRNA vectors (5 μg) with three packaging plasmids: pMDL, VSV-G, and Rev. Forty-eight hours after infection, viral supernatants were collected, filtered through a 0.22 μm filter and added to HCT116 cells. Infected HCT116 cells were selected with puromycin (1 μg/ml) for three consecutive days.

2.4.16 Quantitative real-time RT-PCR (qRT-PCR) analysis

Total RNA was isolated using TRIzol reagent (Life Technologies) following the manufacturer's instructions. cDNA was generated using SuperScript III RT (ThermoFisher Scientific) with random priming. cDNA was subjected to quantitative PCR (q-PCR) using iQ

SYBR green master mix on a CFX96 Real-Time PCR detection System (BioRad) with the primer pairs listed below. The alpha tubulin 1b (TUBA1B) transcript was used for normalization. Experiments were done in technical duplicates and error bars reported as standard deviation. Primers used are as follow; STAT1, F-CTAGTGGAGTGGAAGCGGAG, R-CACCACAAACGAGCTCTGAA; TUBA1B, F-GGCCCCGCCCTAGTGCGTTA, R-GGTGCACTGGTCAGCCAGCTT.

2.4.17 Cell viability assay

Cell viability for HCT116 cells, treated for 48 hours with the indicated largazole concentration or unstimulated (DMSO), was measured using the crystal violet staining method. In short, treated cells were gently washed once with phosphate buffer saline (PBS) and fixed for 20 mins at room temperature with 4% paraformaldehyde under constant rocking. After a single wash with PBS, fixed cells were stained with 0.5% crystal violet (Sigma) in 20% methanol at room temperature for 10 mins. Cells were then thoroughly washed with water and left overnight to dry. Last, 150 µl of developing solution (4:1:1 mix of methanol, ethanol, and water) was added to each well and absorbance was measured at λ = 560 nM.

2.4.18 Flow cytometry analysis

HCT116 cells ($1x10^6$) were treated with vehicle (DMSO) or the indicated dose of largazole for 25 hours. For each cell population analyzed, we washed with ice-cold PBS, treated with trypsin solution, and fixed in cold 70% ethanol overnight. Fixed cells were then washed with ice cold PBS, and incubated in 0.25 mg/ml or RNase (Sigma) for 1 hour at 37°C. Before analysis, cells were stained with 10 ug/ml of propidium iodide (PI) (Sigma) at 4°C for 1 hour. Analysis was performed using a FACScan flow cytometer (Becton-Dickinson). Data obtained from the cell cycle distribution were analyzed using FlowJo version 10.1 (Tree Star). Gaussian

distributions and S-phase polynomial were assigned to each cell population using the Watson pragmatic model (Watson et al., 1987). Starting from samples treated with 9.4 nM largazole dose and above, we specified the range of G1 and G2 peaks in order to gather percentage of cells in each cell cycle phase. Figures were constructed using Microsoft excel version 15.20 or Adobe Illustrator CS6.

Chapter 3: Future directions – classical HDACs, gatekeepers of apoptotic associated *cis*-regulatory elements

3.1 Introduction

Although the extensive impact of HDAC inhibition on chromatin structure and functionality has been widely documented, the specifics about the mode of action remain poorly understood. Through my thesis work I found important clues for understanding the effects of largazole in the colon cancer cell line HCT116. Specifically, I described genome-wide analyses of the effects of increasing largazole-dose cell treatments which challenge general assumptions that chromatin hyperacetylation by HDAC inhibition results in positive stimulation of transcription.

My findings make a significant contribution to the field for four reasons. First, I have developed the most comprehensive analysis to date showing genome-wide dose dependent acetylation changes with the use of a novel High-seq signal calling algorithm (FStitch (Azofeifa et al., 2014)). Second, to my knowledge, this is the first study to demonstrate deacetylation of enhancer-associated H3K27ac as a HDAC inhibition output. Three, I show that HDAC inhibition by largazole treatments results in the systematic emergence of hundreds of enhancer-like elements, suggesting that classical HDACs are responsible for maintaining the inactive state of poised enhancers. Fourth, I found that largazole induces a dose-dependent differential gene expression trend in which low doses primarily induce transcriptional upregulation whereas mid- to high-largazole dose treatments lead to a general mRNA depletion. Together, my data strongly suggest that HDAC inhibitors act through histone deacetylases to modify *cis*-regulatory elements and not just promoter regions as previously thought.

In the following section, I present a discussion about the relevance of my findings, how these observations fit with the current literature, and I also suggest specific investigations that should be considered to better understand and target classical histone deacetylases using small molecules.

3.2 Signal dependent transcription factors and the emergence of enhancer-like elements

Extracellular signals cue cells to differentially express specific gene modules that in turn execute cell programs. Transcription factors and nuclear hormone receptors are mainly responsible for interpreting environmental signals and relaying this information by reorganizing genomic chromatin and consequently modulate gene transcription. Cell-specific transcriptional responses provide a robust system that has been recently shown to be based on the activation of a relatively small number of transcription regulatory factors (Heinz et al., 2010). In this way, signal-dependent transcription factors (SDTFs) act on cell-specific enhancer networks to promote expression of hundreds of genes (Zhang and Glass, 2013). For instance, in macrophages, collective interactions between the lineage-determining factors PU.1 and AP-1 generate a large fraction of the cell-specific sites of open chromatin that give rise to the enhancers responsible for signal-dependent responses to LXR signaling (Heinz et al., 2010). In this system, the synchronous chromatin binding of various SDTFs initiates nucleosome remodeling, including histone acetylation and methylation, at enhancer-like regions that associate with broadly expressed genes.

Based on the large number of new enhancer-like elements established by largazole exposure, I suspect that treated HCT116 cells undergo a similar stimulation that activates SDTFs either at a transcriptional level or via post-translational modifications. Consequently, largazole-activated factors are likely key participants in the formation of enhancer-like elements. As a first

step towards addressing this hypothesis, I searched for mRNAs encoding SDTFs which were significantly upregulated under largazole treatments. Indeed, in addition to the upregulation of several gene members of the AP-1 complex mentioned before, I found the transcriptional activation of other signal dependent factors, including interferon regulatory factors 6 and 7 (IRF6/7), transcription factor AP-2 gamma (TFAP2C), early growth response protein 1 (EGR1), cyclic AMP-dependent transcription factor ATF-3, signal transducers and activators of transcription 1, 2, and 3 (STAT1/2/3) (Fig. 3.1).

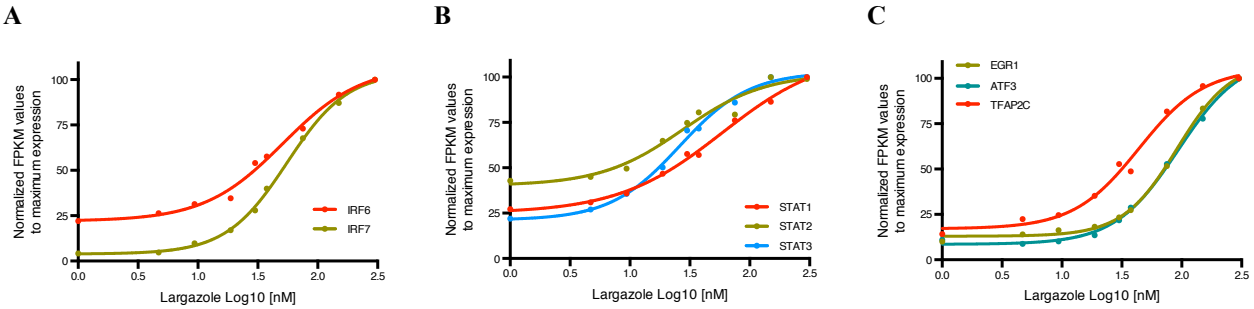


Figure 3.1 Largazole stimulates the transcription of several signal dependent transcription factors
mRNA accumulation levels from HCT116 cells treated for 16 hours with the indicated largazole concentration. Transcript levels are shown as fragments per kilobase of exon per million mapped reads (FPKM) normalized to maximum. A) Interferon regulatory factors 6 and 7. B) Signal transducers and activators of transcription 1, 2, and 3. C) Early growth response protein 1, cyclic AMP-dependent transcription factor 3, and transcription factor AP-2 gamma.

From this list of transcription factors, there are two classes of particular interest: members of the AP-1 complex, since their DNA recognition motif was found heavily enriched among remodeled enhancers, and the signal transducers and activators of transcription (STATs). Several lines of evidence demonstrate that STAT proteins are indispensable in the establishment and suppression of lineage-specific enhancers. Specifically, during differentiation of T helper 1 (Th1) cells, STAT1 and STAT4 work in concert to generate open chromatin and together prime regions to establish a new enhancer landscape (Vahedi et al., 2012). Moreover, interruption of STAT1 or

STAT4 functionality results in deficient recruitment of the acetyl transferase p300 to differentiation-associated enhancer locations.

To test the functional relevance of the transcriptionally activated SDTFs in the establishment of the largazole-phenotype, we have recently treated HCT116 cells with largazole in combination with inhibitors of either the AP-1 complex (T5224) or STAT1/3/5 (SH-4-54). Preliminary cell viability data shows that inactivation of either the AP-1 complex or the STAT1/3/5 proteins results in partial resistance of HCT116 cells to largazole (Fig. 3.2).

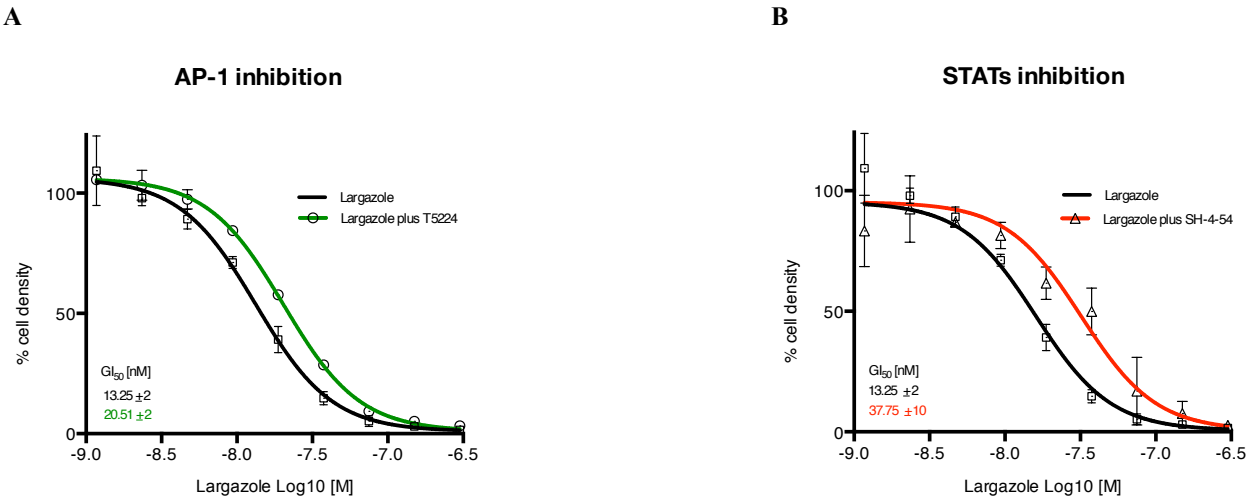


Figure 3.2 Inhibition of AP-1 or STATs function results in partial resistance to largazole-mediated apoptosis in HCT116 cells
Viability of HCT116 cells stimulated for 48 hours under the presence of largazole only (black) or in combination with the inhibitor for A) AP-1 complex (T5224, green), or B) STAT1/3/5 (SH-4-54, red) was determined using crystal violet staining method. Data are presented as mean $\pm$ SD, n=3.

In ongoing investigations, we are also performing individual stable knockdowns of JUN, JUNB, FOS, FOSB, and STAT1, 2, and 3 in HCT116 cells. Cell growth inhibition analysis under incremental largazole dose treatments, showed that STAT1 knockdown cells, display more than

a three-fold increase in cell survival capability relative to control cells, with a half-maximum growth inhibitory concentration (GI_{50}) value of about 287.6 nM (Fig. 3.3).

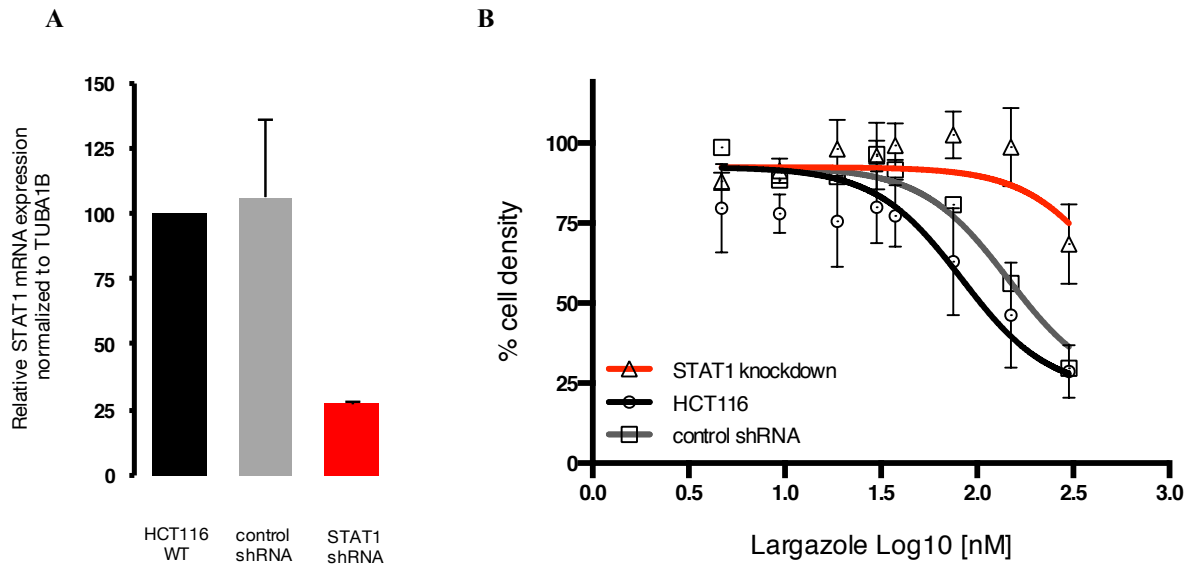


Figure 3.3 Knockdown of STAT1 significantly increases cell viability in HCT116 cells treated with largazole A) Relative mRNA levels of STAT1 gene in WT, control shRNA, and STAT1-knockdown HCT116 cells were determined by real-time qPCR assays (n=2, mean ± SD). B) Viability of 48 hours largazole stimulated HCT116 cells was determined using crystal violet staining method. Data are presented as mean ± SD, n=6.

The higher tolerance to largazole observed in cells with diminished STAT1 protein activity or reduced STAT1 transcript levels implies that SDTFs are contributing to HDAC-inhibition associated cell death and should be further investigated. The SDTFs listed above are excellent candidates for involvement in HDACI-induced cell death and should be examined in a similar manner. Special priority should be given to components of the AP-1 complex (JUN, JUNB, FOS, and FOSB), since AP-1 DNA binding motifs were highly enriched in both largazole-induced decommissioned and emerging enhancers.

3.3 Classical HDACs maintain a repressive state of poised enhancers

Perhaps the most important insight gathered from my thesis investigations gives rise to the idea that classical HDACs maintain poised enhancer regions in a repressed state. Consequently, inhibition of HDACs by largazole allows for acetylation of H3K9/27, which is critical for subsequent enhancer activation. Several previous studies support this hypothesis. In macrophages, the NCoR1/HDAC3 corepressor complex acts as a transcription factor checkpoint. The recruitment of this complex to specific promoter regions with AP-1 DNA binding sites results in the deacetylation of histone tails which is required for transcription. Although a functional NCoR complex is not necessary for macrophage differentiation, NCoR-deficient macrophages upregulate transcripts associated with inflammation, chemotaxis, cell-cycle control, and collagen metabolism (Ogawa et al., 2004).

These observations imply that HDAC complexes could very well be recruited to particular enhancer elements for active histone deacetylation and therefore repression of specific gene transcripts. Indeed, in $CD4^+$ T cells, class I HDACs (HDAC2, and 3) are bound to the intergenic enhancer-like regions at high frequency (Wang et al., 2009b). Moreover, the same genomic locations are co-occupied by several HATs (p300, CBP, PCAF, MOF, and Tip60), suggesting dynamic histone acetylation and deacetylation remodeling at these locations (Wang et al., 2009b). Interestingly, under basal cellular conditions I found high levels of p300 ChIP-seq signal at intergenic locations that will become enhancers upon largazole treatment. In fact, these regions display similar p300 occupancy levels to those of active enhancers in untreated HCT116 cells (Figure 2.15). However, these genomic locations have little or no acetylated H3K9 and –K27 signal.

Together, these observations lead me to ask whether classical HDACs target particular chromatin regions capable of assuming transcriptional enhancer functions and by actively

deacetylating histones maintain repression of these locations. To address this hypothesis, it is necessary to perform two types of high-throughput experiments with HCT116 cells. First, to confirm that class I HDACs are indeed co-occupying the same enhancer regions as p300, individual ChIP-seq experiments with antibodies against HDAC1, 2, and 3 must be conducted. Second, as mentioned above the ultimate proof of enhancer functionality is the production of bi-directional eRNAs; therefore, analyses of nascent transcripts via GRO-seq experiments are required in order to determine the levels of eRNAs at the enhancer locations in question.

3.4 Inactivation of active enhancers by largazole treatment

Classical zinc dependent HDACs were originally discovered based on their ability to remove acetyl groups from histone tails. The observed deacetylation of histone H3K9ac and –K27ac at enhancer regions was unexpected since largazole primarily targets and efficiently inhibits nuclear HDACs. A possible explanation for this incompatible event is that the removal of acetyl groups from enhancer-associated histones is mediated by members of the sirtuin family of deacetylases, which maintain their catalytic activity in the presence of largazole. In vitro studies of sirtuins 1 and 2 demonstrate that these two enzymes can rapidly deacetylate multiple monoacetylated histone tails within a nucleosome structure, including H3K9ac and –K27ac (Hsu et al., 2016). Sirtuins have also been shown to partner with transcription factors and to be recruited to promoter regions (Michishita et al., 2008). Interestingly, in rat cardiomyocites SIRT6 can directly control IGF/Akt signaling at the chromatin level through deacetylation of histone H3 (Sundaresan et al., 2012). In this system, SIRT6 physically interacts with the c-Jun homodimer, is then recruited to proximal promoters containing the AP-1/c-Jun consensus binding site, and inhibits the transcription of IGF signaling-related genes by deacetylating histone H3 at lysine 9. Moreover, SIRT6 knockdowns enhanced the transcription of a reporter

construct driven by endogenous c-Jun and ectopic expression of wild-type SIRT6 but not a catalytically inactive mutant form decreased the transcriptional activity of c-Jun.

These findings, together with my observations that deacetylated enhancer regions showed enrichment for the AP-1 DNA binding site suggest that the coupling of SIRT6 and c-Jun might be responsible for the inactivation of enhancers in HCT116 cells stimulated with largazole. As mentioned above, we are currently exploring the involvement of AP-1 proteins in largazole-mediated cell death. Independent reduction of endogenous c-Jun and SIRT6 proteins with shRNA followed by H3K27ac ChIP-qPRC targeting inactivated enhancers would be sufficient to determine if these proteins are responsible for H3K27ac deacetylation and enhancer decommissioning.

3.5 RNA Pol II pausing at transcription start sites

During the time of my investigation, a similar study was published by the group of Dr. Kim at the University of Texas. Their observations are in agreement and fully support my findings. Using GRO-seq analysis to measure nascent transcripts of human breast carcinoma (BT474) cells stimulated with the pan-specific HDAC inhibitors TSA and SAHA, they found that HDAC inhibition represses gene expression by inducing RNA Pol pausing (Kim et al., 2013) and showed that this event is associated with the redistribution of the bromodomian-containing protein 4 (BRD4), which negotiates transcription elongation (Greer et al., 2015).

Based on these findings, they propose the following model: (1) HDAC inhibition by TSA and SAHA leads to extensive genome hyperacetylation, (2) promoter-bound BRD4 is lost and redistributed mostly along genomic regions devoid of genes, (3) this results in the inability to properly recruit elongation factors at promoters that are necessary for RNA Pol II pause release and engagement of active transcription. In accordance with this model, others have shown that

the activity of HDACs is required for proper localization of BRD4 and complete gene transcription (Hu et al., 2014). Indeed, BRD4 is broadly expressed across many cell lines. For example, in $CD4^+$ T cells it was found associated with 23,518 actively transcribed genes (Zhang et al., 2012), suggesting that the BRD4-dependent RNA Pol II pausing in cells treated with HDAC inhibitors should be a global phenomenon. However, while most of the protein coding regions I analyzed show an increase in pausing of RNA Pol II, there is also a significant group of genes that exhibit a largazole dose-dependent transcription stimulation (Figure 2.17B). For instance, I found that largazole positively influences transcription of FOS at low doses and this trend culminates in a 37-fold induction at the highest drug treatment, from 2.7 to 101.7 FPMKs (Figure 2.16). Interestingly, in Jurkat cells, BRD4 occupies the promoter region of FOS and it is essential for elevated expression levels (Wong et al., 2014). This suggests that in my studies BRD4 is likely being properly recruited to the FOS promoter and mediating full transcription activation, especially in cells with high largazole treatments. This assumption leads me to believe that in addition to BRD4 redistribution there might be other regulatory forces contributing to largazole-induced pausing of RNA Pol II and consequential reduction of associated mRNA levels.

Although the existence of transcriptional enhancers was first reported more than 30 years ago, the study of the epigenetic marks and transcriptional events associated with these genomic regions is a relatively new field. Not surprisingly, there is no general agreement about the direction of information flow between enhancer elements and proximal promoters. However, the concomitant occurrence of three events from mid to high largazole dose treatments hints at a functional relationship between them: (1) inactivation of specific functional enhancers (based on H3K9/27ac, H3K4me1/2, and RNA Pol II occupancy levels), (2) increase in RNA Pol II pausing

at promoters of a subset of genes, (3) and the general depletion of mRNA accumulation levels. Future studies should aim to dissect the influence that largazole-deactivated enhancers have over gene promoters, with special focus on the RNA Pol II pause release transition.

3.6 Conclusions

In summary, my studies revealed a mechanistically novel pathway employed by largazole to regulate apoptosis through the activation of SDTFs in human colon cancer cells. My findings also suggest a more critical involvement of classical HDACs in the regulation of the enhancer landscape than previously thought. Still, one critical observation remains to be explained; largazole induces broad hyperacetylation of both normal and cancer cells (with some exceptions), yet the compound preferentially targets cancer cells for apoptosis. An important clue comes from transcriptome profile analysis revealing that largazole induces minimal mRNA changes in normal fetal colon cells when compared to those from colon cancer. This leads me to ask the following two questions: 1) What are the effects of largazole on the enhancer elements found in normal and differentiated cells? 2) Are the enhancers from non-transformed cells under the repressive control of Zn^{2+} dependent HDACs? Future studies focusing on cell specific enhancer "fingerprints" and the molecular complexes responsible for the functional state of enhancers should provide more insights into the HDAC-inhibition mediated effects of compounds such as largazole and ultimately uncover predictive biomarkers.

Bibliography

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nature reviews. Genetics *13*, 730-31

Allen, M.A., Andrysik, Z., Dengler, V.L., Mellert, H.S., Guarnieri, A., Freeman, J.A., Sullivan, K.D., Galbraith, M.D., Luo, X., Lee Kraus, W., et al. (2014). Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. eLife *2014*, 1–29.

Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., et al. (2007). New Nomenclature for Chromatin-Modifying Enzymes. Cell *131*, 633–636.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Arber, N., Hibshoosh, H., Yasui, W., Sgambato, A., Yamamoto, H., Shapira, I., Rosenman, D., Fabian, I., Weinstein, I.B., Tahara, E., et al. (1999). Abnormalities in the Expression of Cell Cycle-related Proteins in Tumors of the Small Bowel. Cancer Epidemiology Biomarkers and Prevention. *8*, 1101–1105.

Ashkenazi, A. (2002). Targeting death and decoy receptors of the tumour-necrosis factor superfamily. Nature Reviews. Cancer *2*, 420–430.

Augereau, P., and Chambon, P. (1986). The mouse immunoglobulin heavy-chain enhancer: effect on transcription in vitro and binding of proteins present in HeLa and lymphoid B cell extracts. The EMBO Journal *5*, 1791–1797.

Azofeifa, J., Allen, M.A., Lladser, M., and Dowell, R. (2014). FStitch : A fast and simple algorithm for detecting nascent RNA transcripts. Proceedings of the 5th ACM Conference on Bioinformatics Computational Biology, and Health Informatics. 174-183.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: Tools for motif discovery and searching. Nucleic Acids Research *37*.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell *33*, 729–740.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Research *21*, 381–395.

Bao, J., Lu, Z., Joseph, J.J., Carabenciov, D., Dimond, C.C., Pang, L., Samsel, L., McCoy, J.P., Leclerc, J., Nguyen, P., et al. (2010). Characterization of the murine SIRT3 mitochondrial localization sequence and comparison of mitochondrial enrichment and deacetylase activity of long and short SIRT3 isoforms. Journal of Cellular Biochemistry *110*, 238–247.

Barbaric, S., Walker, J., Schmid, a, Svejstrup, J.Q., and Hörz, W. (2001). Increasing the rate of chromatin remodeling and gene activation--a novel role for the histone acetyltransferase Gcn5. The EMBO Journal *20*, 4944–4951.

Barish, G.D., Yu, R.T., Karunasiri, M., Ocampo, C.B., Dixon, J., Benner, C., Dent, A.L., Tangirala, R.K., and Evans, R.M. (2010). Bcl-6 and NF- k B cistromes mediate opposing regulation of the innate immune response. 2760–2765.

Barneda-Zahonero, B., and Parra, M. (2012). Histone deacetylases and cancer. Molecular Oncology *6*, 579–589.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. Cell *129*, 823–837.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell *125*, 315–326.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. Nature Reviews. Drug Discovery *5*, 769–784.

Bolden, J.E., Shi, W., Jankowski, K., Kan, C.-Y., Cluse, L., Martin, B.P., MacKenzie, K.L., Smyth, G.K., and Johnstone, R.W. (2013). HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. Cell Death & Disease *4*, e519.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczyński, B., Riddell, A., and Furlong, E.E.M. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nature Genetics *44*, 148–156.

Bowers, A., West, N., Taunton, J., Schreiber, S.L., Bradner, J.E., and Williams, R.M. (2008). Total synthesis and biological mode of action of largazole: A potent class I histone deacetylase inhibitor. Journal of the American Chemical Society *130*, 11219–11222.

Bowers, A.A., West, N., Newkirk, T.L., Troutman-Youngman, A.E., Schreiber, S.L., Wiest, O., Bradner, J.E., and Williams, R.M. (2009). Synthesis and histone deacetylase inhibitory activity of largazole analogs: Alteration of the zinc-binding domain and macrocyclic scaffold. Organic Letters *11*, 1301–1304.

Brosch, G., Loidl, P., and Graessle, S. (2008). Histone modifications and chromatin dynamics: A focus on filamentous fungi. FEMS Microbiology Reviews *32*, 409–439.

Brown, C.E., Lechner, T., Howe, L., and Workman, J.L. (2000). The many HATs of transcription coactivators. Trends in Biochemical Sciences *25*, 15–19.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843–851.

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327–339.

Butler, L.M., Agus, D.B., Scher, H.I., Higgins, B., Rose, A., Cordon-Cardo, C., Thaler, H.T., Rifkind, R.A., Marks, P.A., and Richon, V.M. (2000). Suberoylanilide Hydroxamic Acid, an Inhibitor of Histone Deacetylase, Suppresses the Growth of Prostate Cancer Cells in Vitro and in Vivo. Cancer Res. *60*, 5165–5170.

Butler, L.M., Zhou, X., Xu, W., Scher, H.I., Rifkind, R.A., Marks, P.A., and Richon, V.M. (2002). The histone deacetylase inhibitor SAHA arrests cancer cell growth , up-regulates thioredoxin-binding protein-2 , and down-regulates thioredoxin. PNAS. *99*, 11700-5

Byers, S.A., Price, J.P., Cooper, J.J., Li, Q., and Price, D.H. (2005). HEXIM2, a HEXIM1-related protein, regulates positive transcription elongation factor b through association with 7SK. Journal of Biological Chemistry *280*, 16360–16367.

Candido, E.P.M., Reeves, R., and Davie, J.R. (1978). Sodium butyrate inhibits histone deacetylation in cultured cells. Cell *14*, 105–113.

Carey, M., Lin, Y.S., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature *345*, 361–364.

Carroll, J.S., Meyer, C. a, Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., et al. (2006). Genome-wide analysis of estrogen receptor binding sites. Nature Genetics *38*, 1289–1297.

Chavanas, S., Adoue, V., Méchin, M.C., Ying, S., Dong, S., Duplan, H., Charveron, M., Takahara, H., Serre, G., and Simon, M. (2008). Long-range enhancer associated with chromatin looping allows AP-1 regulation of the peptidylarginine deiminase 3 gene in differentiated keratinocyte. PLoS ONE *3*, 1–11.

Chen, G., Fernandez, J., Mische, S., and Courey, A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor Groucho in Drosophila development. Genes and Development *13*, 2218–2230.

Cheng, B., and Price, D.H. (2007). Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. Journal of Biological Chemistry *282*, 21901–21912.

Chi, P., Allis, C.D., and Wang, G.G. (2010). Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nature Reviews. Cancer *10*, 457–469.

Choi, O.R., and Engel, J.D. (1986). A 3' enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. Nature *323*, 731–734.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annual Review of Biochemistry *78*, 273–304.

Cole, K.E., Dowling, D.P., Boone, M.A., Phillips, A.J., and Christianson, D.W. (2011). Structural basis of the antiproliferative activity of largazole, a depsipeptide inhibitor of the histone deacetylases. Journal of the American Chemical Society *133*, 12474–12477.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nature Genetics *46*, 1311–1320.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931–21936.

Davidson, I., Fromental, C., Augereau, P., Wildeman, a, Zenke, M., and Chambon, P. (1986). Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. Nature *323*, 544–548.

Davie, J.R. (2003). Inhibition of histone deacetylase activity by butyrate. The Journal of Nutrition *133*, 2485S – 2493S.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233–1244.

Deqing Hu, Alexander S. Garruss, Xin Gao, Marc A. Morgan, Malcolm Cook, E.R., and Smith, A.S. (2013). The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse

embryonic stem cells. Nat Struct Mol Biol. *29*, 1093–1097.

Eckner, R., Ewen, M.E., Newsome, D., Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., Decaprio, J.A., Lawrence, J.B., and Livingston, D.M. (1994). Molecular cloning and functional analysis of the adenovirus E1A-assoclated 300-kD protein.(p300) reveals a protein with properties oI a transcriptional adaptor. 869–884.

Eissenberg, J.C., and Elgin, S.C. (1991). Boundary functions in the control of gene expression. Trends in Genetics : TIG *7*, 335–340.

Enderle, D., Beisel, C., Stadler, M.B., Gerstung, M., Athri, P., and Paro, R. (2011). Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. Genome Research *21*, 216–226.

Euskirchen, G.M., Auerbach, R.K., Davidov, E., Gianoulis, T.A., Zhong, G., Rozowsky, J., Bhardwaj, N., Gerstein, M.B., and Snyder, M. (2011). Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. PLoS Genetics *7*.

Falkenberg, K.J., and Johnstone, R.W. (2014). Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. Nature Reviews. Drug Discovery *13*, 673–691.

Federico, M., and Bagella, L. (2011). Histone deacetylase inhibitors in the treatment of hematological malignancies and solid tumors. J Biomed Biotechnol *2011*, 475641.

Finzer, P., Kuntzen, C., Soto, U., zur Hausen, H., and Rösl, F. (2001). Inhibitors of histone deacetylase arrest cell cycle and induce apoptosis in cervical carcinoma cells circumventing human papillomavirus oncogene expression. Oncogene *20*, 4768–4776.

Forneris, F., Binda, C., Vanoni, M.A., Battaglioli, E., and Mattevi, A. (2005). Human histone demethylase LSD1 reads the histone code. Journal of Biological Chemistry *280*, 41360–41365.

Frietze, S., Wang, R., Yao, L., Tak, Y., Ye, Z., Gaddis, M., Witt, H., Farnham, P.J., and Jin, V.X. (2012). Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. Genome Biology *13*, R52.

Galloni, M., Gyurkovics, H., Schedl, P., and Karch, F. (1993). The bluetail transposon: evidence for independent cis-regulatory domains and domain boundaries in the bithorax complex. The EMBO Journal *12*, 1087–1097.

Giese, K., Kingsley, C., Kirshner, J.R., and Grosschedl, R. (1995). Assembly and function of a TCRα enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. Genes and Development *9*, 995–1008.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable

precise gene regulation. Cell *143*, 540–551.

Gilmour, D.S., and Fan, R. (2009). Detecting transcriptionally engaged RNA polymerase in eukaryotic cells with permanganate genomic footprinting. Methods *48*, 368–374.

Goodman, R.H., and Smolik, S. (2000). CBP / p300 in cell growth , transformation , and development. Genes & Development *14*, 1553–1577.

Greer, C.B., Tanaka, Y., Kim, Y.J., Xie, P., Zhang, M.Q., Park, I.H., and Kim, T.H. (2015). Histone Deacetylases Positively Regulate Transcription through the Elongation Machinery. Cell Reports *13*, 1444–1455.

Gregoretti, I. V., Lee, Y.M., and Goodson, H. V. (2004). Molecular evolution of the histone deacetylase family: Functional implications of phylogenetic analysis. Journal of Molecular Biology *338*, 17–31.

Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F.F., Zamir, I., et al. (1998). Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. Nature *391*, 815–818.

Grozinger, C.M., and Schreiber, S.L. (2002). Deacetylase enzymes: Biological functions and the use of small-molecule inhibitors. Chemistry and Biology *9*, 3–16.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. Cell *130*, 77–88.

Gui, C.-Y., Ngo, L., Xu, W.S., Richon, V.M., and Marks, P. a (2004). Histone deacetylase (HDAC) inhibitor activation of p21WAF1 involves changes in promoter-associated proteins, including HDAC1. Proceedings of the National Academy of Sciences of the United States of America *101*, 1241–1246.

Guillemette, B., Drogaris, P., Lin, H.H.S., Armstrong, H., Hiragami-Hamada, K., Imhof, A., Bonneil, ??ric, Thibault, P., Verreault, A., and Festenstein, R.J. (2011). H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. PLoS Genetics *7*.

Guillou, E., Ibarra, A., Coulon, V., Casado-Vela, J., Rico, D., Casal, I., Schwob, E., Losada, A., and M??ndez, J. (2010). Cohesin organizes chromatin loops at DNA replication factories. Genes and Development *24*, 2812–2822.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome Biology *8*, R24.

Gyurkovics, H., Gausz, J., Kummer, J., and Karch, F. (1990). A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation. The EMBO Journal *9*, 2579–2585.

Haberland, M., Montgomery, R.L., and Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. Nature Reviews. Genetics *10*, 32–42.

Hagstrom, K., Muller, M., and Schedl, P. (1996). Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the Drosophila bithorax complex. Genes and Development *10*, 3202–3215.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013). Enhancer transcripts mark active estrogen receptor binding sites Enhancer transcripts mark active estrogen receptor binding sites. 1210–1223.

Hall, J.A., Dominy, J.E., Lee, Y., and Puigserver, P. (2013). The sirtuin family's role in aging and age-associated pathologies. Journal of Clinical Investigation *123*, 973–979.

Hallson, G., Hollebakken, R.E., Li, T., Syrzycka, M., Kim, I., Cotsworth, S., Fitzpatrick, K.A., Sinclair, D.A.R., and Honda, B.M. (2012). dSet1 is the main H3K4 Di- and tri-methyltransferase throughout Drosophila development. Genetics *190*, 91–100.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: The next generation. Cell *144*, 646–674.

Haumaitre, C., Lenoir, O., and Scharfmann, R. (2008). Histone deacetylase inhibitors modify pancreatic cell fate determination and amplify endocrine progenitors. Molecular and Cellular Biology *28*, 6373–6383.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature Genetics *39*, 311–318.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell *38*, 576–589.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol *16*, 144–154.

Herz, H.M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y. hei, Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by trithorax-related, the drosophila homolog of mammalian MLL3/MLL4. Genes and Development *26*, 2604–2620.

Hesse, J.E., Nickol, J.M., Lieber, M.R., and Felsenfeld, G. (1986). Regulated gene expression in transfected primary chicken erythrocytes. Proc Natl Acad Sci U S A *83*, 4312–4316.

Hong, J., and Luesch, H. (2012). Largazole: From discovery to broad-spectrum therapy. Natural Product Reports *29*, 449.

Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. Nature Reviews Molecular Cell Biology *13*, 225–238.

Hsin, J.-P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes & Development *26*, 2119–2137.

Hsu, W.W., Wu, B., and Liu, W.R. (2016). Sirtuins 1 and 2 are Universal Histone Deacetylases. ACS Chemical Biology acschembio.5b00886.

Hu, D., Gao, X., Morgan, M. a, Herz, H.-M., Smith, E.R., and Shilatifard, A. (2013). The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Molecular and Cellular Biology *33*, 4745–4754.

Hu, X., Lu, X., Liu, R., Ai, N., Cao, Z., Li, Y., Liu, J., Yu, B., Liu, K., Wang, H., et al. (2014). Histone Cross-talk Connects Protein Phosphatase 1 (PP1) and Histone Deacetylase (HDAC) Pathways to Regulate the Functional Transition of Bromodomain-containing 4 (BRD4) for Inducible Gene Expression. Journal of Biological Chemistry *289*, 23154–23167.

Huang, B., Yang, X.-D., Zhou, M.-M., Ozato, K., and Chen, L.-F. (2009a). Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. Molecular and Cellular Biology *29*, 1375–1387.

Huang, D.W., Lempicki, R. a, and Sherman, B.T. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols *4*, 44–57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009c). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research *37*, 1–13.

Huang, J., Sengupta, R., Espejo, A.B., Lee, M.G., Dorsey, J.A., Richter, M., Opravil, S., Shiekhattar, R., Bedford, M.T., Jenuwein, T., et al. (2007). p53 is regulated by the lysine demethylase LSD1. Nature *449*, 105–108.

Hubbert, C., Guardiola, A., Shao, R., Kawaguchi, Y., Ito, A., Nixon, A., Yoshida, M., Wang, X.-F., and Yao, T.-P. (2002). HDAC6 is a microtubule-associated deacetylase. Nature *417*, 455–458.

Jamaladdin, S., Kelly, R.D.W., O'Regan, L., Dovey, O.M., Hodson, G.E., Millard, C.J., Portolano, N., Fry, A.M., Schwabe, J.W.R., and Cowley, S.M. (2014). Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America *111*, 9840–9845.

Ji, R.R., Siemers, N.O., Lei, M., Schweizer, L., and Bruccoleri, R.E. (2011). SDRS-An algorithm for analyzing large-scale dose-response data. Bioinformatics *27*, 2921–2923.

John, S., Howe, L., Tafrov, S.T., Grant, P.A., Sternglanz, R., and Workman, J.L. (2000). The Something About Silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF. Genes & Development 1196–1208.

Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Developmental Biology *339*, 225–229.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., et al. (2013). Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Molecular Cell *51*, 310–325.

Kanno, T., Kanno, Y., LeRoy, G., Campos, E., Sun, H.-W., Brooks, S.R., Vahedi, G., Heightman, T.D., Garcia, B. a, Reinberg, D., et al. (2014). BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nature Structural & Molecular Biology *21*, 1047–1057.

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genomics *13*, 424.

Kawakami, E., Tokunaga, A., Ozawa, M., Sakamoto, R., and Yoshida, N. (2015). The histone demethylase Fbxl11/Kdm2a plays an essential role in embryonic development by repressing cell-cycle regulators. Mechanisms of Development *135*, 31–42.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, a. D. (2002). The Human Genome Browser at UCSC.

Kim, Y.Z. (2014). Altered histone modifications in gliomas. Brain Tumor Research and Treatment *2*, 7–21.

Kim, T.K., and Maniatis, T. (1997). The mechanism of transcriptional synergy of an in vitro assembled interferon-beta enhanceosome. Molecular Cell *1*, 119–129.

Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D. a, Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Kim, Y.J., Greer, C.B., Cecchini, K.R., Harris, L.N., Tuck, D.P., and Kim, T.H. (2013). HDAC

inhibitors induce transcriptional repression of high copy number genes in breast cancer through elongation blockade. Oncogene *32*, 2828–2835.

Kim, Y.W., Lee, S., Yun, J., and Kim, A. (2015). Chromatin looping and eRNA transcription precede the transcriptional activation of gene in the β-globin locus. Bioscience Reports *35*, 1–8.

Kollias, G., Hurst, J., deBoer, E., and Grosveld, F. (1987). The human beta-globin gene contains a downstream developmental specific enhancer. Nucleic Acids Res *15*, 5739–5747.

Kooistra, S.M., and Helin, K. (2012). Molecular mechanisms and potential functions of histone demethylases. Nature Publishing Group *13*, 297–311.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nature Reviews. Molecular Cell Biology *12*, 283–294.

Kumar, R., Li, D., Müller, S., and Knapp, S. (2016). Epigenomic regulation of oncogenesis by chromatin remodeling. Oncogene 1–14.

Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. Annual Review of Genetics *47*, 483–508.

Lagouge, M., Argmann, C., Gerhart-Hines, Z., Meziane, H., Lerin, C., Daussin, F., Messadeq, N., Milne, J., Lambert, P., Elliott, P., et al. (2006). Resveratrol Improves Mitochondrial Function and Protects against Metabolic Disease by Activating SIRT1 and PGC. Cell *127*, 1109–1122.

Lai, F., and Shiekhattar, R. (2014). Enhancer RNAs: The new molecules of transcription. Current Opinion in Genetics and Development *25*, 38–42.

Lam, M.T.Y., Li, W., Rosenfeld, M.G., and Glass, C.K. (2014). Enhancer RNAs and regulated transcriptional programs. Trends in Biochemical Sciences *39*, 170–182.

Lane, A.A., and Chabner, B.A. (2009). Histone deacetylase inhibitors in cancer therapy. J Clin Oncol *27*, 5459–5468.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology *10*, R25.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends in Genetics *32*, 42–56.

Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. Nature Reviews. Molecular Cell Biology *8*, 284–295.

Lee, T.I., and Young, R. a (2000). Transcription of Eukaryotic Protein-Coding Genes. Annual Review of Genetics 77–137.

LeRoy, G., Rickards, B., and Flint, S.J. (2008). The Double Bromodomain Proteins Brd2 and Brd3 Couple Histone Acetylation to Transcription. Molecular Cell *30*, 51–60.

Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Human Molecular Genetics *12*, 1725–1735.

Levine, M. (2010). Transcriptional enhancers in animal development and evolution. Current Biology *20*, R754–R763.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.

Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A.K., and Rosenfeld, M.G. (2013a). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. Cell *155*, 1581–1595.

Liu, Y., Salvador, L.A., Byeon, S., Ying, Y., Kwan, J.C., Law, B.K., Hong, J., and Luesch, H. (2010). Anticolon cancer activity of largazole, a marine-derived tunable histone deacetylase inhibitor. The Journal of Pharmacology and Experimental Therapeutics *335*, 351–361.

Liu, Y., Wang, Z., Wang, J., Lam, W., Kwong, S., Li, F., Friedman, S.L., Zhou, S., Ren, Q., Xu, Z., et al. (2013b). A histone deacetylase inhibitor, largazole, decreases liver fibrosis and angiogenesis by inhibiting transforming growth factor-β and vascular endothelial growth factor signalling. Liver International : Official Journal of the International Association for the Study of the Liver *33*, 504–515.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320–334.

Luger, K., Mäder, a W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Malik, S., and Roeder, R.G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. Nature Reviews. Genetics *11*, 761–772.

Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R., and Pazdur, R. (2007). FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. The

Oncologist *12*, 1247–1252.

Marks, P. a (2010). Histone deacetylase inhibitors: a chemical genetics approach to understanding cellular functions. Biochimica et Biophysica Acta *1799*, 717–725.

Marks, P.A., Richon, V.M., and Rifkind, R.A. (2000). Histone Deacetylase Inhibitors: Inducers of Differentiation or Apoptosis of Transformed Cells. JNCI Journal of the National Cancer Institute *92*, 1210–1216.

Mathew, O.P., Ranganna, K., and Yatsu, F.M. (2010). Butyrate , an HDAC inhibitor , stimulates interplay between different posttranslational modifications of histone H3 and differently alters G1-specific cell cycle proteins in vascular smooth muscle cells. Biomedicine et Pharmacotherapy *64*, 733–740.

Matthews, K. (1992). DNA Looping. Microbiology and Molecular Biology Reviews *56*, 123–136.

Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. Nature Biotechnology *31*, 1133–1136.

Mercola, M., Goverman, J., Mirell, C., and Calame, K. (1985). Immunoglobulin heavy-chain enhancer requires one or more tissue-specific factors. Science *227*, 266–270.

Metzger, E., Wissmann, M., Yin, N., Müller, J.M., Schneider, R., Peters, A.H.F.M., Günther, T., Buettner, R., and Schüle, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436–439.

Michishita, E., McCord, R.A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T.L.A., Barrett, J.C., et al. (2008). SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. Nature *452*, 492–496.

Miller, T., Krogan, N.J., Dover, J., Tempst, P., Johnston, M., Greenblatt, J.F., Sherman, D.R., Voskuil, M., Schnappinger, D., Harrell, M.I., et al. (2001). COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *98*.

Minucci, S., and Pelicci, P.G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nature Reviews Cancer *6*, 38–51.

Mohan, M., Herz, H.-M., Smith, E.R., Zhang, Y., Jackson, J., Washburn, M.P., Florens, L., Eissenberg, J.C., and Shilatifard, a. (2011). The COMPASS Family of H3K4 Methylases in Drosophila. Molecular and Cellular Biology *31*, 4310–4318.

Moon, K.J., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Molecular Cell *19*, 523–534.

Mottamal, M., Zheng, S., Huang, T.L., and Wang, G. (2015). Histone deacetylase inhibitors in clinical studies as templates for new anticancer agents. Molecules *20*, 3898–3941.

Müller, B.M., Jana, L., Kasajima, A., Lehmann, A., Prinzler, J., Budczies, J., Winzer, K.-J., Dietel, M., Weichert, W., and Denkert, C. (2013). Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer--overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. BMC Cancer *13*, 215.

Nagarajan, S., Hossan, T., Alawi, M., Najafova, Z., Indenbirken, D., Bedi, U., Taipaleenm??ki, H., Ben-Batalla, I., Scheller, M., Loges, S., et al. (2014). Bromodomain Protein BRD4 Is Required for Estrogen Receptor-Dependent Enhancer Activation and Gene Transcription. Cell Reports *8*, 460–469.

Nakagawa, M., Oda, Y., Eguchi, T., Aishima, S.-I., Yao, T., Hosoi, F., Basaki, Y., Ono, M., Kuwano, M., Tanaka, M., et al. (2007). Expression profile of class I histone deacetylases in human cancer tissues. Oncology Reports *18*, 769–774.

Narlikar, G.J., Sundaramoorthy, R., and Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. Cell *154*, 490–503.

Ocker, M. (2010). Deacetylase inhibitors - focus on non-histone targets and effects. World Journal of Biological Chemistry *1*, 55–61.

Ogawa, S., Lozach, J., Jepsen, K., Sawka-Verhelle, D., Perissi, V., Sasik, R., Rose, D.W., Johnson, R.S., Rosenfeld, M.G., and Glass, C.K. (2004). A nuclear receptor corepressor transcriptional checkpoint controlling activator protein 1-dependent gene networks required for macrophage activation. Proceedings of the National Academy of Sciences of the United States of America *101*, 14461–14466.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281–300.

Parthun, M.R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. Oncogene *26*, 5319–5328.

Pekowska, A., Benoukraf, T., Zacarias-Cabeza, J., Belhocine, M., Koch, F., Holota, H., Imbert, J., Andrau, J.-C., Ferrier, P., and Spicuglia, S. (2011). H3K4 tri-methylation provides an epigenetic signature of active enhancers. The EMBO Journal *30*, 4198–4210.

Peterlin, B.M., and Price, D.H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. Molecular Cell *23*, 297–305.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. Nature *322*, 697–701.

Qiu, H., Hu, C., Yoon, S., Natarajan, K., Swanson, M.J., and Hinnebusch, A.G. (2004). An array

of coactivators is required for optimal recruitment of TATA binding protein and RNA polymerase II by promoter-bound Gcn4p. Molecular and Cellular Biology *24*, 4104–4117.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rada-Iglesias, A., Enroth, S., Ameur, A., Koch, C.M., Clelland, G.K., Respuela-Alonso, P., Wilcox, S., Dovey, O.M., Ellis, P.D., Langford, C.F., et al. (2007). Butyrate mediates decrease of histone acetylation centered on transcription start sites and down-regulation of associated genes. Genome Research *17*, 708–719.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. a, Flynn, R. a, and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rafehi, H., Balcerczyk, A., Lunke, S., Kaspi, A., Ziemann, M., Harikrishnan, K.N., Okabe, J., Khurana, I., Ooi, J., Khan, A.W., et al. (2014). Vascular histone deacetylation by pharmacological HDAC inhibition. Genome Research *24*, 1271–1284.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). C-Myc regulates transcriptional pause release. Cell *141*, 432–445.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

Riester, D., Hildmann, C., Gr??newald, S., Beckers, T., and Schwienhorst, A. (2007). Factors affecting the substrate specificity of histone deacetylases. Biochemical and Biophysical Research Communications *357*, 439–445.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W.W.M.P., Wilm, M., Aasland, R., and Stewart, A.F. (2002). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. EMBO Journal *20*, 7137–7148.

Rosa, S., and Shaw, P. (2013). Insights into Chromatin Structure and Dynamics in Plants. 1378–1410.

Roth, S.Y., Denu, J.M., and Allis, D. (2001). Histone Acetyltransferases. Annu. Rev. Biochem *70*, 81–120.

de Ruijter, A.J.M., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B.P. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. The Biochemical Journal *370*, 737–749.

Sadakierska-Chudy, A., and Filip, M. (2014). A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin

Regulation by ncRNAs. Neurotoxicity Research *27*, 172–197.

Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T., and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). Mammalian Genome *15*, 23–34.

Schaaf, C.A., Misulovin, Z., Gause, M., Koenig, A., Gohara, D.W., Watson, A., and Dorsett, D. (2013). Cohesin and Polycomb Proteins Functionally Interact to Control Transcription at Silenced and Active Genes. PLoS Genetics *9*.

Schaukowitch, K., Joo, J.Y., Liu, X., Watts, J.K., Martinez, C., and Kim, T.K. (2014). Enhancer RNA facilitates NELF release from immediate early genes. Molecular Cell *56*, 29–42.

Scholer, H.R., and Gruss, P. (1985). Cell type-specific transcriptional enhancement in vitro requires the presence of trans-acting factors. *4*, 3005–3013.

Sen, R., and Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell *46*, 705–716.

Seumois, G., Chavez, L., Gerasimova, A., Lienhard, M., Omran, N., Kalinke, L., Vedanayagam, M., Ganesan, A.P. V, Chawla, A., Djukanović, R., et al. (2014). Epigenomic analysis of primary human T cells reveals enhancers associated with TH2 memory cell differentiation and asthma susceptibility. Nature Immunology *15*, 777–788.

Shaulian, E., and Karin, M. (2001). AP-1 in cell proliferation and survival. Oncogene *20*, 2390–2400.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat Cell Biol *4*, E131–E136.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941–953.

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Molecular Cell *19*, 857–864.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual Review of Biochemistry *81*, 65–95.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nature Reviews. Genetics *15*, 272–286.

Simon, A., and Wolfgang, H. (2010). Differential expression analysis for sequence count data. Genome Biology *11*, R106.

Smith, B.C., and Denu, J.M. (2009). Chemical mechanisms of histone lysine and arginine

modifications. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms *1789*, 45–57.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nature Reviews Genetics *13*, 613–626.

Stasevich, T.J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., McNally, J.G., et al. (2014). Regulation of RNA polymerase II activation by histone acetylation in single living cells. Nature *516*, 272–275.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Su, W., Porter, S., Kustu, S., and Echols, H. (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. Proceedings of the National Academy of Sciences of the United States of America *87*, 5504–5508.

Sun, W., Zhou, X., Zheng, J., Lu, M., Nie, J., and Yang, X. (2012). Histone acetyltransferases and deacetylases : molecular and clinical implications to gastrointestinal carcinogenesis. Acta Biochimica et Biophysica Sinica *44*, 80–91.

Sundaresan, N.R., Vasudevan, P., Zhong, L., Kim, G., Samant, S., Parekh, V., Pillai, V.B., Ravindra, P. V, Gupta, M., Jeevanandam, V., et al. (2012). The sirtuin SIRT6 blocks IGF-Akt signaling and development of cardiac hypertrophy by targeting c-Jun. Nature Medicine *18*, 1643–1650.

Svechnikova, I., Almqvist, P.M., and Ekstrom, T.J. (2008). HDAC inhibitors effectively induce cell type-specific differentiation in human glioblastoma cell lines of different origin. Int J Oncol *32*, 821–827.

Taatjes, D.J. (2010). The human Mediator complex: A versatile, genome-wide regulator of transcription. Trends in Biochemical Sciences *35*, 315–322.

Taft, R.J., Pheasant, M., and Mattick, J.S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. BioEssays *29*, 288–299.

Takahashi, Y. -h., Westfield, G.H., Oleskie, a. N., Trievel, R.C., Shilatifard, a., and Skiniotis, G. (2011). Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. Annual review of biochemistry . *81*, 65-95

Tang, Z., Chen, W.-Y., Shimada, M., Nguyen, U.T.T., Kim, J., Sun, X.-J., Sengoku, T., McGinty, R.K., Fernandez, J.P., Muir, T.W., et al. (2013). SET1 and p300 Act Synergistically, through Coupled Histone Modifications, in Transcriptional Activation by p53. Cell *154*, 297–310.

Taori, K., Paul, V.J., and Luesch, H. (2008). Structure and activity of largazole, a potent

antiproliferative agent from the Floridian marine cyanobacterium Symploca sp. Journal of the American Chemical Society *130*, 1806–1807.

Thomas, M.C., and Chiang, C.-M. (2006). The general transcription machinery and general cofactors. Critical Reviews in Biochemistry and Molecular Biology *41*, 105–178.

Thurman, R., Rynes, E., Humbert, R., Vierstra, J., Maurano, M., Haugen, E., Sheffield, N., Stergachis, A., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Tie, F., Banerjee, R., Stratton, C. a, Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M.O., Scacheri, P.C., and Harte, P.J. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development (Cambridge, England) *136*, 3131–3141.

Tie, F., Banerjee, R., Saiakhova, A.R., Howard, B., Monteith, K.E., Scacheri, P.C., Cosgrove, M.S., and Harte, P.J. (2014). Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. Development (Cambridge, England) *141*, 1129–1139.

Tie, F., Banerjee, R., Fu, C., Stratton, C.A., Fang, M., and Harte, P.J. (2015). Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. 1–10.

Tjeertes, J. V, Miller, K.M., and Jackson, S.P. (2009). Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. The EMBO Journal *28*, 1878–1889.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and De Laat, W. (2002). Looping and interaction between hypersensitive sites in the active ??-globin locus. Molecular Cell *10*, 1453–1465.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trudel, M., and Costantini, F. (1987). A 3’ enhancer contributes to the stage-specific expression of the human beta-globin gene. Genes & Development *1*, 954–961.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811–816.

Tsukiyama, T. (2002). The in vivo functions of ATP-dependent chromatin-remodelling factors. Nature Reviews. Molecular Cell Biology *3*, 422–429.

Turner, N.C., Ro, J., André, F., Loi, S., Verma, S., Iwata, H., Harbeck, N., Loibl, S., Huang Bartlett, C., Zhang, K., et al. (2015). Palbociclib in Hormone-Receptor-Positive Advanced Breast

Cancer. The New England Journal of Medicine 1–11.

Ungermannova, D., Parker, S.J., Nasveschuk, C.G., Wang, W., Quade, B., Zhang, G., Kuchta, R.D., Phillips, A.J., and Liu, X. (2012). Largazole and its derivatives selectively inhibit ubiquitin activating enzyme (E1). PLoS ONE *7*.

Vahedi, G., Takahashi, H., Nakayamada, S., Sun, H.W., Sartorelli, V., Kanno, Y., and OShea, J.J. (2012). STATs shape the active enhancer landscape of T cell populations. Cell *151*, 981–993.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1__γ are associated with transcription elongation through mammalian chromatin. Molecular Cell *19*, 381–391.

Ververis, K., Hiong, A., Karagiannis, T.C., and Licciardi, P. V. (2013). Histone deacetylase inhibitors (HDACIS): Multitargeted anticancer agents. Biologics: Targets and Therapy *7*, 47–60.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K. a, et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390–394.

Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Research *22*, 1798–1812.

Wang, P., Lin, C., Smith, E.R., Guo, H., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M., et al. (2009a). Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Molecular and Cellular Biology *29*, 6074–6085.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nature Genetics *40*, 897–903.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009b). Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. Cell *138*, 1019–1031.

Watson, J. V, Chambers, S.H., and Smith, P.J. (1987). A pragmatic approach to the analysis of DNA histograms with a definable G1 peak. Cytometry *8*, 1–8.

Weintraub, H., and Groudine, M. (1976). Chromosomal Subunits in Active Genes Have an Altered Conformation. Science *193*.

West, A.C., and Johnstone, R.W. (2014). New and emerging HDAC inhibitors for cancer treatment. J Clin Invest. *124*, 30–39.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221–225.

Wingender, E., Schoeps, T., Haubrock, M., and D??nitz, J. (2015). TFClass: A classification of human transcription factors and their rodent orthologs. Nucleic Acids Research *43*, D97–D102.

Wissmann, M., Yin, N., Müller, J.M., Greschik, H., Fodor, B.D., Jenuwein, T., Vogler, C., Schneider, R., Günther, T., Buettner, R., et al. (2007). Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. Nature Cell Biology *9*, 347–353.

Wolf, D., Rodova, M., Miska, E.A., Calvet, J.P., and Kouzarides, T. (2002). Acetylation of ??-catenin by CREB-binding protein (CBP). Journal of Biological Chemistry *277*, 25562–25567.

Wong, M.M., Byun, J.S., Sacta, M., Jin, Q., Baek, S., and Gardner, K. (2014). Promoter-bound p300 complexes facilitate post-mitotic transmission of transcriptional memory. PLoS ONE *9*.

Xu, M., Nie, L., Kim, S.H., and Sun, X.H. (2003). STAT5-induced Id-1 transcription involves recruitment of HDAC1 and deacetylation of C/EBP?? EMBO Journal *22*, 893–904.

Xu, W., Edmondson, D.G., Roth, S.Y., and Iol, M.O.L.C.E.L.L.B. (1998). Mammalian GCN5 and P / CAF Acetyltransferases Have Homologous Amino-Terminal Domains Important for Recognition of Nucleosomal Substrates. *18*, 5659–5669.

Xu, W., Ngo, L., Perez, G., Dokmanovic, M., and Marks, P. a (2006). Intrinsic apoptotic and thioredoxin pathways in human prostate cancer cell response to histone deacetylase inhibitor. Proc Natl Acad Sci U S A *103*, 15540–15545.

Xu, W.S., Parmigiani, R.B., and Marks, P. (2007). Histone deacetylase inhibitors: molecular mechanisms of action. Oncogene *26*, 5541–5552.

Yamaguchi, Y., Shibata, H., and Handa, H. (2013). Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms *1829*, 98–104.

Yamane, K., Toumazou, C., Tsukada, Y. ichi, Erdjument-Bromage, H., Tempst, P., Wong, J., and Zhang, Y. (2006). JHDM2A, a JmjC-Containing H3K9 Demethylase, Facilitates Transcription Activation by Androgen Receptor. Cell *125*, 483–495.

Yang, X., and Gregoire, S. (2005). Class II Histone Deacetylases : from Sequence to Function , Regulation , and Clinical Implication MINIREVIEW Class II Histone Deacetylases : from Sequence to Function , Regulation , and Clinical Implication. Molecular and Cellular Biology *25*, 2873–2884.

Yang, X.-J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria

and yeast to mice and men. Nature Reviews. Molecular Cell Biology *9*, 206–218.

Yang, X.J., Ogryzko, V. V, Nishikawa, J., Howard, B.H., and Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature *382*, 319–324.

Yang, Z., Yik, J.H.N., Chen, R., He, N., Moon, K.J., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Molecular Cell *19*, 535–545.

Ying, Y., Taori, K., Kim, H., Hong, J., and Luesch, H. (2008). Total synthesis and molecular target of largazole, a histone deacetylase inhibitor. Journal of the American Chemical Society *130*, 8455–8459.

Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. Nature Publishing Group *21*, 564–578.

Zeng, X., Yin, B., Hu, Z., Liao, C., Liu, J., Li, S., Li, Z., Nicklaus, M.C., Zhou, G., and Jiang, S. (2010). Total synthesis and biological evaluation of largazole and derivatives with promising selectivity for cancers cells. Organic Letters *12*, 1368–1371.

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Research *21*, 1273–1283.

Zhang, D.X., and Glass, C.K. (2013). Towards an understanding of cell-specific functions of signaldependent transcription factors. Journal of Molecular Endocrinology *51*.

Zhang, K., Williams, K., Huang, L., Yau, P., Siino, J., Bradbury, E.M., Jones, P., Minch, M., and Burlingame, A. (2002). Histone Acetylation and Deacetylation Identification of Acetylation and Methylation Sites of HeLa Histone H4 by Mass Spectrometry. Molecular & Cellular Proteomics *1*, 500–508.

Zhang, W., Prakash, C., Sum, C., Gong, Y., Li, Y., Kwok, J.J.T., Thiessen, N., Pettersson, S., Jones, S.J.M., Knapp, S., et al. (2012). Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells. Journal of Biological Chemistry *287*, 43137–43155.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*.

Zhao, Y., Lu, J., Sun, H., Chen, X., Huang, W., Tao, D., and Huang, B. (2005). Histone acetylation regulates both transcription initiation and elongation of hsp22 gene in Drosophila. Biochemical and Biophysical Research Communications *326*, 811–816.

Zhou, J., Wang, X., He, K., Charron, J.B.F., Elling, A.A., and Deng, X.W. (2010). Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in arabidopsis reveals correlation

between multiple histone marks and gene expression. Plant Molecular Biology *72*, 585–595.

Zhu, Y., Sun, L., Chen, Z., Whitaker, J.W., Wang, T., and Wang, W. (2013). Predicting enhancer transcription and activity from chromatin modifications. Nucleic Acids Research *41*, 10032–10043.

Zupkovitz, G., Grausenburger, R., Brunmeir, R., Senese, S., Tischler, J., Jurkin, J., Rembold, M., Meunier, D., Egger, G., Lagger, S., et al. (2010). The cyclin-dependent kinase inhibitor p21 is a crucial target for histone deacetylase 1 as a regulator of cellular proliferation. Mol Cell Biol *30*, 1171–1181.