2014

Chromatin Control of the Antiviral Response to Influenza

Jessica Sook Yuin Ho

CHROMATIN CONTROL OF THE ANTIVIRAL RESPONSE TO INFLUENZA

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Jessica Sook Yuin Ho

June 2014

CHROMATIN CONTROL OF THE ANTIVIRAL RESPONSE TO INFLUENZA

Jessica Sook Yuin Ho, Ph.D.

The Rockefeller University 2014

A key development for our understanding of the mechanisms that control gene expression has been the finding that the histones recruit proteins with effector functions to chromatin. This is mediated primarily by post-translational modifications that occur on the histone N-terminal domains ("tails"). Single or combinations of histone tail modifications serve as scaffolds for protein complexes controlling transcription or co-transcriptional processes, thus impacting gene expression. Histone tail modifications are regulated by multiple, often overlapping pathways in the cell, and as such, present an important regulatory "node" through which the cell is able to integrate and respond to environmental signals. However, a consequence of this is that any artificial or naturally occurring molecule that "mimics" the histone tails has the potential to strongly impact gene function and the cell's response to the environment.

Indeed, we were able to identify a novel pathway exploited by the influenza virus to directly dampen the host transcriptional response. The non-structural protein 1 (NS1) of the Influenza virus contains a histone H3-like sequence that is able to bind to and disrupt the activity of the human PAF1 transcription elongation complex (PAF1C). Loss of PAF1C function leads to an impaired antiviral response and increased influenza viral replication. Genome-wide binding analyses indicate that PAF1 is inducibly recruited to anti-viral and inflammatory genes during infection, and that its presence coincides with

the recruitment of RNA polymerase II (Pol II) and the expression of target genes. Altogether, our findings imply that exploiting histone mimicry could be a general strategy for pathogens to subvert or co-opt host-processes for their own benefit. Our studies also strongly suggest that proper regulation of transcription elongation by PAF1C is an important rate-limiting step in the transcriptional response to pathogens.

To my Mother, Father, Brother and Brian

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor Dr. Sasha Tarakhovsky for his support and guidance through these years. His creativity and unique perspective on science has been a constant inspiration for me. He has taught me how to think independently as a scientist, and most importantly, how to be scientifically vigorous. Every day in his lab has been an incredible learning experience, and I feel privileged to have had that opportunity to study under him.

I also thank the members of my faculty advisory committee, Dr. Charles Rice, Dr. David Allis, for all their time and their helpful suggestions through the years. They have always enthusiastically supported me and encouraged me. I also thank Dr. Peter Palese for taking the time of his busy schedule to serve on my committee as my external member.

I would like to extend my deepest gratitude to Dr. Ivan Marazzi, who took me under his wing, trained and guided me all the way through graduate school. He was not only a wonderful scientific mentor, but also a good friend, who cheered me on when I was feeling discouraged. He also tried (and failed miserably) to expand my knowledge about movies and music, but I applaud him for his patience to that respect.

I thank our collaborators, Dr. Adolfo García-Sastre, Dr. Ana Fernandez-Sesma, Dr. Balaji Manicassamy, Dr. Randy Albrecht and Richard Cadagan for all their time and advice with regards to conducting virus experiments. They were always generous in sharing reagents (and viruses), and were never ever too busy to teach me new techniques or give me advice about science. Chris Seibert also generously gave us antibodies for our experiments with virus.

Dr. Scott Dewell was an amazing collaborator, who helped us work through the complicated entity that is high-throughput sequencing. He was amazingly patient with all the naïve questions about sequencing that I had, and was always ready to help when I need it.

Dr. Robert Roeder and Dr. Jaehoon Kim were our collaborators on the PAF1 project, and without whom the *in vitro* transcriptional assays with PAF1 would not have existed. I would like to thank them for all their help and contributions to this project. They also generously provided reagents for our pull-down assays.

I would also like to thank all past and present members of the Tarakhovsky lab for their selfless help and advice through graduate school. Special thanks go to Dr. Uwe Schaefer, who helped me through many of the experiments that involved ChIP-seq. He has always been extremely generous with sharing his reagents, and was always there in times of need. I would also like to thank Dr. Ryan Kim, who helped me greatly with the

bioinformatics analysis for the PAF1 ChIP-seq experiments, and Jenieve Guevarra, who helped me with my experiments, especially towards the final legs of my thesis. I would also like to acknowledge Angela Santana, our lab manager, who has always helped me out of tight spots in lab, and been a wonderful friend to me. I thank also Raffi Cohn and Sophie Huang for helping me with preparation and initial analysis of the PAF1 ChIP sequencing libraries.

The past and present members of the Rockefeller Genomics Resource Center have helped me with my (many) high throughput experiments with their excellent expertise.

Many thanks also go to the members of the Dean's Office whom have helped me at numerous occasions, and gave me the opportunity to perform this work.

I would also like to acknowledge Dr. Edward Ruby and Dr. Margaret McFall-Ngai, whose deep passions for science (and host-microbe interactions) were decidedly the sources of inspiration for me to attempt taking on graduate school.

Last, but not least, I would like to thank all of my family and friends, who have encouraged me and cheered me on all through these years – especially to my parents, Mr. and Mrs. Ho Ah Chai and my brother, Ian who bore all the ups and downs of graduate school with me and were always there when I needed them. I also want to acknowledge Uncle Nee and Aunty Molly, who have always worried after me. Finally, I would also like to thank my partner Brian. His unwaveringly love, patience and support for me every step of the way has made all this possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Selective Responses to Infectious Agents

The inflammatory response represents the first line of defense against invading microorganisms. This response relies on the activities of a multitude of germline encoded pattern recognition receptors (PRRs) that are able to detect specific pathogen-derived macromolecules and components, commonly referred to as 'pathogen-associated molecular patterns (PAMPs) (Baccala et al., 2009; Kawai and Akira, 2007). Some PRRs may also be involved in sensing damage-associated molecular patterns (DAMPs), which are endogenous indicators of cellular damage (Matzinger, 1994; Seong and Matzinger, 2004).

Known PRRS include the Toll-like receptors (TLRs), retinoic acid-inducible gene I (RIG-I)-like receptors, nucleotide oligomerization domain (NOD)-like receptors, as well as C-type lectin receptors(Akira et al., 2006; Kawai and Akira, 2007). Upon recognition of their cognate ligands, these receptors activate signaling cascades that lead to the activation of various sequence-specific transcription factors, including nuclear factor kappa-B (NF-κB) and the interferon regulatory factors (IRFs). Together, these transcription factors coordinate the transcriptional response against pathogens. This typically involves the expression of pro-inflammatory cytokines, such as interleukin-6 (IL-6), tumor necrosis factors-alpha (TNFα) or the type I interferons (IFNα/β). In turn, these cytokines function to induce expression of anti-microbial genes and to recruit effectors cells of the innate and adaptive immune system, all in all facilitating pathogen clearance.

While the inflammatory response is critical for the clearance of pathogenic intruders, inappropriate induction or prolongation of the inflammatory response may result in damage to host tissues and threaten organism survival. As such, the cells have evolved numerous strategies to control both the kinetics and magnitude of the inflammatory response during infection. Aside from this, the cell must also develop strategies to defend against antagonism by pathogen-derived effector molecules. I focus here on regulation of the inflammatory response at the level of gene expression and the chromatin-based mechanisms that controls it.

1.2 Overview of Chromatin Structure

The genomes of eukaryotic cells are maintained as a stable nucleoprotein-DNA complex called chromatin. Chromatin is organized in a hierarchical manner, with the nucleosome as its fundamental repeating subunit (Kornberg, 1974; Olins and Olins, 1974). Nucleosome bound tracts of DNA account for the first organizational level of chromatin and appear by electron microscopy as “beads on a string” (where nucleosomes are the “beads” and the intervening DNA is the “string”). Poly-nucleosome bound tracts of DNA can then be further folded into higher ordered structures with increasing DNA packing densities (Figure 1.1). Linker histones, which bear little structural resemblance to the core histones, are believed to facilitate compaction by mediating inter-nucleosomal interactions(Thomas, 1999; Vignali and Workman, 1998).

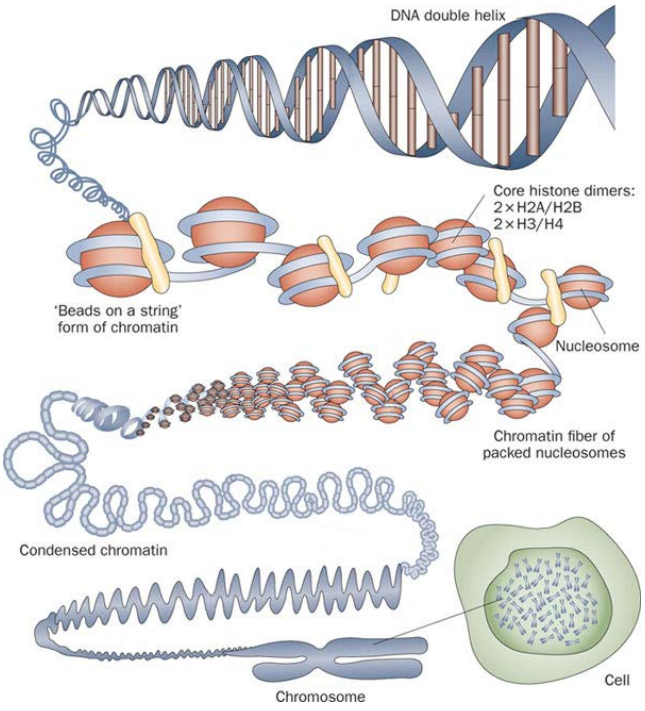


Figure 1.1: Overview of chromatin organization Genomes of eukaryotic cells are organized into chromatin, which is composed of repeating units of nucleosomes. Nucleosomes are connected to each other by short stretches of linker DNA, forming the first organizational level of chromatin ("beads-on-a-string"). Nucleosome bound tracts of DNA may fold in higher order structures to form condensed chromatin fibers. (Image adapted from Tonna et al., 2010)

At the heart of chromatin, lies the nucleosome core particle. Each particle comprises of approximately 147 base pairs of DNA wrapped in a 1.65 super helical turns around an octamer of four 'core' histone proteins (H2A, H2B, H3 and H4)(Luger et al., 1997). Octamers are formed from a central H3:H4 tetramer (formed from two H3:H4 dimers) that is flanked by two H2A: H2B dimers (Davey et al., 2002; Luger et al., 1997). The histone proteins are relatively similar in structure, featuring a C-terminal "histone-fold" domain as well as a short, unstructured N-terminal "tail" domain (Figure 1.2) (Davey et al., 2002; Luger et al., 1997). The histone fold domains of the histone proteins

are responsible for mediating both histone-histone as well as histone-DNA interactions, and form the bulk of the nucleosome core structure (Hacques et al., 1990; Luger et al., 1997). On the other hand, the N-terminal tail domains of the histone proteins are thought to extrude DNA bound nucleosome core particle. These exposed N-terminal "tail" domains are also rich in basic residues, and are subject to a number of post-translational modifications (See also section 1.4)

Packaging DNA into nucleosomes reduces DNA accessibility to cellular factors, and thus has the potential to hinder cellular processes dependent on DNA (including transcription (Huang and Bonner, 1962; Laybourn and Kadonaga, 1991; Morse, 1989; Orphanides and Reinberg, 2000) , replication (Kelly et al., 2010) and DNA repair (Fernandez-Capetillo et al., 2002)). Ability to regulate nucleosome structure, and functionally couple chromatin states to cellular states is thus crucial for the eukaryotic cell's transcriptional response to the environment (including to infection). In the cell, chromatin structure is directly controlled by the activities of two classes of regulators: ATP-dependent chromatin-remodeling complexes and histone-modifying enzymes.

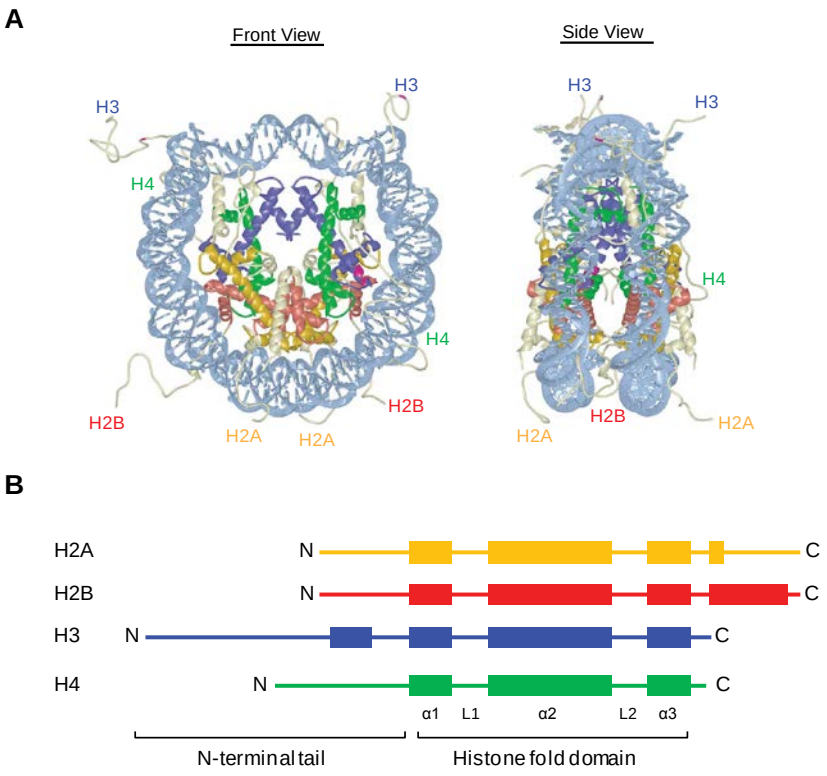


Figure 1.2: Structure of the nucleosome core particle and the histone proteins. A. Front (left panel) and side (right panel) views of the nucleosome core particle. The nucleosome comprises of an octamer of the four core histone proteins, H2A (Yellow), H2B (red), H3 (blue) and H4 (green). Octamers are formed from a central H3:H4 tetramer (formed from two H3:H4 dimers) that is flanked by two H2A: H2B dimers. Approximately 147 base pairs of DNA are wrapped around the nucleosome in 1.65 helical turns. (Images adapted from Luger, 2003) **B.** Linear organization of the core histones. The core histones are structurally similar to each other, comprising of a C-terminal histone-fold domain, as well as an N-terminal unstructured domain ('tail'). Each histone fold domain is made of 3 alpha helices (α1, α2 and α3), separated by loops(L1,L2). In chromatin, the histone-fold domains form the bulk of the nucleosome structure, whereas intrinsically disorder tail domains protrude out of the nucleosome. (Images adapted from Dutnall and Ramakrishnan, 1997)

1.3 ATP-dependent chromatin remodeling complexes

Chromatin remodeling complexes are multi-subunit complexes that utilize ATP hydrolysis to alter histone-DNA interactions. Some of these complexes may function to

bring about the exchange or eviction of nucleosomes, while others are required for nucleosomal repositioning/sliding along the DNA (Mohrmann and Verrijzer, 2005). Ultimately, these activities change the overall accessibility of nucleosome bound DNA for processes like transcription and replication.

Chromatin remodeling complexes are divided into 4 families, based on their ATPase subunit: SWI/SNF, ISWI, CHD/Mi-2 and Ino80 (Becker and Horz, 2002; Eberharter and Becker, 2004; Mohrmann and Verrijzer, 2005). These complexes display distinct remodeling activities, and may work to promote either activating or repressive chromatin environments. For instance, the ATP-utilizing chromatin assembly and remodeling factor (ACF) which is a member of the ISWI family, functions by generating regularly spaced nucleosomes across the DNA, thus restricting DNA accessibility and repressing DNA-dependent processes (Corona et al., 2002; Ito et al., 1999; Shogren-Knaak et al., 2006a). On the other hand, the SWI/SNF family of complexes has been shown to create nucleosome free regions at the promoters and transcription start sites of target genes, enhancing transcription factor binding and the recruitment of the transcriptional machinery (Agalioti et al., 2000; Imbalzano et al., 1994; Ramirez-Carrozzi et al., 2009).

1.4 Histone Modifications and Chromatin dynamics

Histone modifying enzymes constitute the second major class of chromatin regulators. As mentioned previously, the core histone proteins are subject to multiple post-translational modifications. The majority of the PTMs occur in the unstructured, basic N-terminal tail domains. Histone modifications include lysine methylation,

acetylation, sumoylation and ubiquitination; arginine methylation; and serine and threonine phosphorylation (Figure 1.3)(Kouzarides, 2007). These modifications are largely regulated by two sets of enzymatic complexes: those that catalyze the addition of a chemical group/protein to the histone tail substrate ("writers") and those that catalyze its removal ("erasers"). The activities of these enzyme complexes are typically site-specific and many of them are found in conjunction with signaling molecules and larger regulatory complexes.

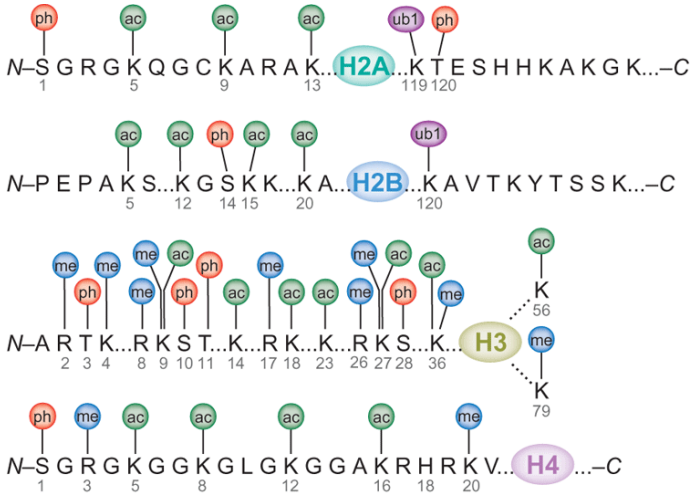


Figure 1.3: Post-translational modification of the core histone proteins Important histone modifications include phosphorylation (ph), acetylation (ac), methylation (me) and ubiquitylation (ub1). Most of the known histone modifications occur on the N-terminal unstructured domains of the histone proteins, although those occurring in the globular histone-fold domains (e.g. H3 at K56 and K79) have been reported to occur. (Images adapted from Bhaumik et al., 2007)

Histone modifications may affect nucleosome and chromatin structure in three non-mutually exclusive ways. First, modifications may change the change nucleosome charge density, and/or interfere with histone-DNA interactions. This could lead to the loss of nucleosome stability or a change in nucleosome mobility. In support of this, force

spectroscopy studies report that hyper-acetylated nucleosomes are less stable than hypo-acetylated nucleosomes (Brower-Toland et al., 2005).

Secondly, histone modification may impact inter-nucleosomal interactions and inhibit the formation of higher order chromatin structures. Acetylation of lysine 16 in histone H4 (H4K16), in particular, was found to prevent condensation of nucleosome arrays into chromatin fibers *in vitro* (Shogren-Knaak et al., 2006b; Zhou et al., 2007) . This could be in part explained by observations that residues 16 to 20 of the N-terminal tail of histone H4 interacts with two acidic-patches in the histone-fold domain in histone H2A of the adjacent nucleosome (Luger et al., 1997). H4K16 acetylation could function either to neutralize electrostatic interactions (Shogren-Knaak and Peterson, 2006) between H4 N-terminal tail and H2A, or to occlude H4 N-terminal tail-H2A interactions.

Finally, the presence of specific histone modifications along with their valency, can be recognized or "read" by dedicated proteins through specialized protein domains (Ruthenburg et al., 2007b; Taverna et al., 2007). Examples of such domains include bromodomains that recognize histone acetyl-lysines; and PHD domains, Tudor domains and chromodomains that bind to methylated lysines or arginines (Table 1.1). Importantly, some reader proteins may also carry two or more of these domains, allowing them to recognize multiple modifications simultaneously(Ruthenburg et al., 2007a). Thus, single and combinatorial patterns of histones could have very different functional outcomes on chromatin, depending on the proteins that recognize them. This forms the basis of the 'histone code' hypothesis (Strahl and Allis, 2000; Turner, 2000).

Table 1.1: Recognition modules of known histone modifications. This table lists examples of known chromatin associated "reader"- modules and the histone post-translational modifications they are known to bind to. Examples of cellular proteins that bear these domains are also indicated. (Adpated from (Taverna et al., 2007) and (Yun et al., 2011))

Reader Module		Position	Known Proteins
Bromodomain		Various	Rsc4, PB1, Brdt, Brd2, Brd3, Brd4
14-3-3		H3S10ph	14-3-3 proteins
BRCT		H2AXS139	MDC1
PHD		H3K4me0	BHC80, AIRE
		H3K4me3	BPTF, TAF3, RAG2, PHF8
		H3K9me3	SMCX
WD40		H3R2	WDR5
Royal Family	Chromodomain	H3K4me1/2/3	CHD1
		H3K9me2/3	HP1
		H3K23me	MPP8
		H3K27me2/3	PC, MPP8
		H3K36me2/3	Eaf3, MSL3, MRG15
	Tudor	H3K4me	JMJD2A, JMJD2C
		H3K9me2/3	TDRD7, UHRF1
		H3K79me2	53BP1
		H4K20	53BP1/Crb2, PHF20
		H3R17	TDRD3
		H4R3	TDRD3
	MBT	H4K20me1/2	PHF20L1
		H3K4me1	PHF20L1
		H3K9me1/2	SFMBT
Ankyrin Repeats		H3K9me2/3	G9a/GLP

1.5 Chromatin dynamics and the Inflammatory Response

As discussed before, chromatin structure imposes obstacles on transcription. As such, chromatin structure and its dynamic regulation can play an important role in determining transcriptional activation or repression of important inflammatory genes

during homeostasis. A prime example of this would be chromatin-based regulation of the IFNβ gene, which is an important cytokine induced in most somatic cell types upon viral infection. Under steady state conditions, the IFNβ gene is kept repressed by the presence of a nucleosome directly positioned at, and obscuring the TATA box at the promoter (Lomvardas and Thanos, 2001). However, during stimulation, acetylation of histones H3 at lysine 9 and lysine 14 and H4 at lysine 8 within this nucleosome (Agalioti et al., 2002) results in nucleosomal remodeling by the SWI/SNF complex and the recruitment of TFIID to the promoter (Agalioti et al., 2000; Lomvardas and Thanos, 2001; Panne et al., 2007; Parekh and Maniatis, 1999). These activities ultimately expose the DNA to the cell's transcriptional machinery, allowing for transcription of the gene to occur.

In addition to regulation during the activation phase, the chromatin environment may serve to limit the overall activity of a given gene. Indeed, di-methylation of histone H3 at lysine 9 (H3K9me2) within nucleosomes bound to the IFNβ promoter was shown to correlate with its transcriptional output (Fang et al., 2012). Reduction of H3K9me2 abundance at the IFNβ promoter, through genetic ablation or pharmacological inhibition of the G9a/GLP methyltransferase complex, correlated with an increased and more rapid expression of IFNβ during gene activation. Levels of H3K9me2 at the IFNβ promoter were thus suggested to be a determinant of cell-type specific differences in IFNβ expression (Fang et al., 2012).

Beyond influencing transcription activation and repression, chromatin structure may influence the activation kinetics of different subsets of genes during the inflammatory response. Indeed, several groups have showed that the inducible recruitment of some transcription factors to their target genes upon immune stimulation is highly dependent

on the pre-existing chromatin state on the target gene. For example, Saccani et al. showed that NF-κB associates with its target genes with variable kinetics during LPS stimulation of macrophages (Saccani et al., 2001, 2002). Certain genes were bound by NF-κB and transcribed immediately upon NF-κB translocation to the nucleus ('primary response genes'), whereas others were bound and transcribed with significantly delayed kinetics ('secondary response genes')(Saccani et al., 2001). This was attributed to the presence of a nucleosome barrier at the secondary response genes, which could be overcome if the cells were pre-treated/primed with IFNγ prior to LPS stimulation.

In support of this, a subsequent study showed that genes activated in LPS-stimulated macrophages showed variable dependence on SWI/SNF chromatin remodeling complex (Ramirez-Carrozzi et al., 2009; Ramirez-Carrozzi et al., 2006). In fact, the majority of LPS-induced primary response genes in this study were activated in a SWI/SNF independent manner, whereas SWI/SNF dependence was observed for the secondary response genes (Ramirez-Carrozzi et al., 2009). Importantly, the differential dependence on SWI/SNF activity correlated strongly with nuclease accessibility at the promoters of target genes; Genes that were SWI/SNF dependent exhibited low nuclease accessibility prior to LPS stimulation, suggesting that nucleosomal organization at the gene was indeed important for determining the kinetics of gene activation (Ramirez-Carrozzi et al., 2006).

1.6 Pathogenic subversion/co-opting of host chromatin processes

Given the importance of chromatin dynamics and histone modifications in the regulation of the immune response to pathogens, it is no wonder that many pathogens

have developed strategies that specifically target host chromatin processes. For instance, the *Shigella flexneri* effector protein OspF migrates to the nucleus during infection, where it specifically targets the ERK (extracellular signal regulated kinases) and p38 MAPKs (mitogen-activated protein kinases) for de-phosphorylation (Arbibe et al., 2007). This inhibits MAPK mediated histone H3 serine 10 phosphorylation, resulting in inhibited transcriptional activation of the inflammatory genes and to the benefit of the bacteria(Arbibe et al., 2007).

Chromatin dynamics has also been shown to play a major role in the life cycles of many viruses. In particular, DNA viruses, which replicate in the nucleus, must contend with the host-mediated chromatinization of their genomes during infection. Several DNA and RNA viruses thus evolved mechanisms that co-opt several host chromatin processes to regulate viral gene expression. For instance, during herpes simplex virus (HSV) infection, viral VP16 protein is required to recruit histone acetyltransferases and chromatin remodeling complexes to the chromatinized viral genome, allowing for the activation of viral genes (Herrera and Triezenberg, 2004). In the absence of VP16, nucleosomes assembled on the viral genome restrict DNA accessibility and repress the viral gene transcription (Herrera and Triezenberg, 2004).

Viruses that integrate into the genome, such as the retroviruses have also evolved strategies to use host chromatin processes for their benefit. For one, heterochromatic silencing of integrated human immunodeficiency virus (HIV) has been suggested to facilitate maintenance latent reservoirs of virally infected cells in the host and evade immune clearance (Pierson et al., 2000). In fact, reactivation of HIV from latency involves chromatin remodeling and the generation of open chromatin at the integrated

viral genome. This is mediated by the HIV protein Trans-activator of transcription, TAT, which recruits histone acetyltransferases (HATs), including CBP/p300 and p/CAF (Deng et al., 2000; Deng et al., 2001; Kiernan et al., 1999), to the viral promoter. Hyperacetylation of nucleosomes by HATs at the viral promoter facilitates chromatin remodeling, and the formation of a permissive chromatin environment for transcription (Lusic et al., 2003). Finally, TAT also recruits the positive transcription elongation factor-b (PTEF-b), which phosphorylates the C-terminal domain of Pol II to promote transcription elongation through the viral genome (Yang et al., 2005).

1.7 Beyond the histone code: Other means to subvert host chromatin processes

The specialized effector domain present in 'readers' recognize short amino acid sequences. Structural studies have also proven that the modularity of this recognition can be artificially modified and that histone tail modifications do not necessarily have to be recognized in the context of the entire histone tail (Li et al., 2007). Rather, the histone tails could be envisioned as being composed of multiple overlapping short linear motifs, with each motif (and its modification state) functioning as a discrete unit of information for histone- and chromatin-bound proteins (Fischle et al., 2003).

An interesting implication of this hypothesis is that non-histone proteins that bear similar motifs might be subject to the same regulation and protein-protein interactions afforded to the histones. This raises an interesting consideration related to host-pathogen interactions: *Can pathogens take advantage of the histone code by supplementing their own proteome with histone-derived motifs?*

Several histone-like sequences in non-histone proteins have already been well characterized in the recent years. One of these proteins is the methyltransferase G9a, which is responsible for histone H3 di-methylation at lysine 9. G9a bears a 163-ARKT-166 motif that strongly resembles the 7-ARKS-10 motif of its target H3 target residue (Figure 1.4A). Consistent with this, G9a was found to auto-methylate itself on lysine 165 (Chin et al., 2007; Sampath et al., 2007), and this methylation was required for G9a to form a co-repressor complex with a separate H3K9me2 reader and chromodomain-containing protein, HP1γ (Sampath et al., 2007).

In addition to this, the H3-like sequence in G9a is also conserved in its homologue and hetero-dimerization partner GLP (Tachibana et al., 2005) (Figure 1.4A), even though the two proteins share relatively poor primary sequence conservation in their N-terminal domains. Like G9a, the H3-like sequence of GLP is subject to auto-methylation. GLP methylation is thought to create a binding site for another chromodomain containing protein M phase phosphoprotein 8, MPP8 (Chang et al., 2011). Interestingly, the chromodomain of MPP8 also interacts with methylated DNMT3A (a *de novo* DNA methyltransferase), which turns out to also carry a similar histone H3K9-like sequence that is recognized and methylated by the G9a/GLP complex (Figure 1.4A). MPP8 dimers binding both methylated G9a/GLP complex and methylated DNMT3A were thus proposed to functionally link two transcriptionally repressive events (i.e. H3K9 methylation and DNA methylation activities) to the same complex (Figure 1.4B) (Chang et al., 2011).

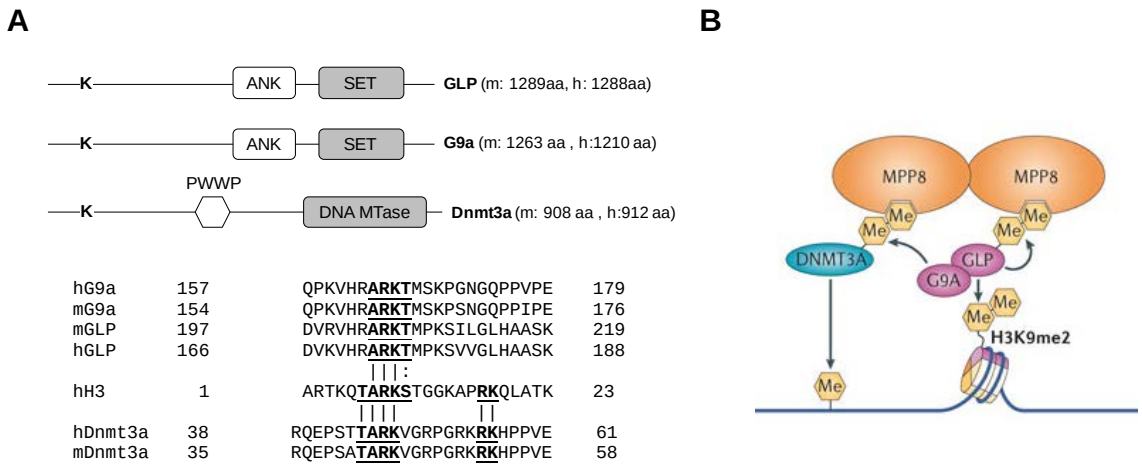


Figure 1.4: Histone mimicry in G9a, GLP and DNMT3a proteins. A. Domain architecture and aligned sequences for mouse (m) or human (h) G9a, GLP and Dnmt3a proteins against histone H3. **B.** Model for GLP/G9a complex interactions with Dnmt3a through MPP8 protein. MPP8 homo-dimers bind to automethylated G9a/GLP complex and methylated Dnmt3a protein. This could potentially couple G9a/GLP mediated H3K9 di-methylation with Dnmt3a mediated *de novo* DNA methylation. ANK: Ankyrin repeats, SET: Su(var)3-9, E(z) Trithorax domain; DNA Mtase: DNA methyltransferase; PWWP: PWWP protein interaction domain. (Images adapted from Sampath et al., 2007; Badeaux and Shi, 2013)

1.8 Motif mimicry is prevalent amongst pathogens

Motif mimicry is a common tactic used by pathogens to subvert or co-opt host cell processes (Davey et al., 2011)(see also Table 1.2). In fact, the putative histone-derived motifs are highly reminiscent of a class of compact, non-globular protein interaction interfaces known as short linear motifs (SLiMs) (Davey et al., 2012; Diella et al., 2008; Van Roey et al., 2013). SLiMS are found widely throughout the genome, and serve many regulatory functions, such as directing ligand binding, serving as sites for post-translational modification and mediating complex assembly. Like the histone –tail motifs, multiple SLiMs can overlap each other, or be used in a cooperative fashion. The ability of

some SLiMs to be post-translationally modified also allows them to alternate between different functional states that can impact downstream protein-protein interactions (Davey et al., 2012; Diella et al., 2008; Van Roey et al., 2013). Thus, the prevalence of these motifs in the regulation of cell function, coupled with their intrinsic properties, has made them ideal targets for pathogens(Davey et al., 2011).

Perhaps, the most well described example of motif mimicry in pathogens would be within the adenovirus E1A protein(Pelka et al., 2008). The E1A protein contains a collection of independent protein binding motifs that allow it to interact with a diverse array of host proteins. These include host CtBP (co-repressor C-terminal binding protein) through a short PxDLS motif (where x is any amino acid) (Boyd et al., 1993; Schaeper et al., 1995); cell cycle regulator pRb through a LxCxE motif(Carvalho et al., 1995); BS69 co-repressor through a PxLxP motifs(Ansieau and Leutz, 2002); and CBP/p300 through a FxD/ExxxL motif (O'Connor et al., 1999). E1A mediated interactions with these proteins allow the virus to target cell cycle- and growth related gene expression, thus enhancing viral replication. Further examples of viral mimicry of host derived motifs and their functional outcomes are displayed in Table 1.2.

Bacteria pathogens have also been shown to mimic host short motifs, although these do not seem to occur to the extent as in viruses, which are obligate intracellular parasites. For instance, the *Vibrio cholera* heat labile toxin (Cholera toxin, CT) contains a C-terminal ER rentention signal KDEL (Sixma et al., 1991). Upon entry into the ER, ER-resident chaperones and enzymes facilitate toxin activation (Tsai et al., 2001). The activated toxin is then able to enhance activation of cellular adenyl cyclase, resulting in a cascade of events that lead to an increased excretion of chloride ions and water from the

affected cells (Sharp and Hynie, 1971). Production of CT is thought to facilitate host colonization, and may also be a mechanism for the bacterium to generate cAMP (via cellular adenyl cyclase) as an energy source.

Interestingly, some viruses have developed effector proteins that hijack the functions of cellular motifs. An example of this occurs in the Papillomaviruses (PV). The PV targets several cellular proteins, including E6AP, interferon regulatory factor-3 (IRF3), the notch co-activator MAMLI, all of which contain acidic leucine (L)-rich sequences comprising an LxxLL motif(Zanier et al., 2013). These interactions are mediated by a unique fold within the PV E6 oncoprotein, and are essential for the virus to suppress the host immune response and induce oncogenesis (Zanier et al., 2013). Loss of the LxxLL binding site in the E6 protein results in the loss of transformation and degradation activities of the E6 protein (Zanier et al., 2013). Altogether, these studies highlight the important role of motif biology in the cell, and why they represent attractive targets for pathogens to manipulate.

Table 1.2: Examples of pathogen mimicy of host-derived short motifs
Motifs are presented in single letter amino acid code; X refers to any amino acid.[‡] E1A Conserved region 1; [†]E1A conserved region 3; $: Protein C-terminus. Amino acids indicated within square braces can be substituted for each other at that position. Adapted from Davey et al., 2011

Host Target	Viral Protein	Virus	Motif	Outcome	Ref.
AP-1	Nef	HIV	ExxxLL	CD4 downregulation; enhanced viral infectivity	Craig et al., 1998
Calcineurin	p12	HTLV1	SPxLxLT	Inhibition of NFAT-calcineurin interactions; Transcriptional repression	Kim et al., 2003
CtBP	E1A (CR1)‡	Adenovirus	PxDLS	Loss of CTBP1 acitivity; Enhanced transformation of cells	Schaeper et al., 1995
CtBP	E1A (CR3)†	Adenovirus	RxxTG	Loss of CTBP1 acitivity; Enhanced transformation of cells	Bruton et al., 2008
Farnesyltransferase	HDAg-L	HDV	Cxxx$	HADg-L Farnesylation; Required for viral biogenesis	Glenn et al., 1992
HCF	VP16	HSV	EHxY	Activation of viral immediate early genes; HCF-1/OCT-1/VP16 complex formation	Lu et al., 1998
JAK	LMP1	EBV	PxxPxP	Activation of NF-KB signaling; Repression of host apoptotic pathways	Gires et al., 1999
NEDD4	VP40	Ebola	PPxY	Ubiquitylation of VP40; Role in viral budding	Harty et al., 2000
Oligosaccharyltransferase	E1	HCV	Nx[ST]	Glycosylation of E1; Role in protein folding and viral entry	Meunier et al., 1999
p300/CBP	E1A	Adenovirus	FxDxxxL	Enhanced expression of viral genes	Ferreon et al., 2009
PDZ domain(s)	E6	HPV	x[ST]xV$	Targetted degradation of host tumor suppressors MAGI-I and SAP97/Dlg	Zhang et al., 2007
PDZ domain(s)	NS1	Influenza A virus	x[ST]xV$	Interactions with host PDZ domains-containing proteins	Obenauer et al., 2006
RB	E1A	Adenovirus	LxxLYD	Displacement of E2F proteins from Rb	Liu and Marmorstein, 2007
SIAH1	ORF45	KSHV	PxAxV	Degradation of ORF45; Putative role in viral re-activation	Abada et al., 2008
TR	E1A	Adenovirus	LxxLIxxxL	Dysregulation of thyroid hormone receptor function	Meng et al., 2005
TRADD	LMP1	EBV	YYD$	Activation of NF-KB signaling; Repression of host apoptotic pathways	Izumi and Kieff, 1997
TRAF2	LMP1	EBV	PxQxT	Activation of NF-KB signaling; Repression of host apoptotic pathways	Ye et al., 1999
Tsg101	Gag	HIV	PTAP	Recruitment of Tsg101 to endosome; Role in viral budding	Pornillos et al., 2002

1.9 Why Histone mimicry?

There are several reasons as to why histone motif mimicry might be a particularly successful strategy for pathogens:

Evolutionary plasticity

For one, these motifs are short, with the capacity to encode a functional interaction interface within three to ten amino acids (four in the case of G9a, GLP and DNMT3A (Chang et al., 2011; Chin et al., 2007; Sampath et al., 2007)). Generation of a functional motif *de novo* in unrelated protein is thus not likely to require more than a few mutations. In addition, the short sequence length also allows pathogens to utilize multiple motifs at any one time. For pathogens, this is an especially important consideration, given that their evolutionary space is constrained by their small genomes.

Modular functionality.

Secondly, as part of the histone protein, these motifs play especially important roles in regulating chromatin structure and gene expression. Incorporation of individual and/or combinations of such histone motifs into a pathogen-derived protein could be sufficient for the pathogen to gain control over entire chromatin regulatory pathways in the cell. Specifically, pathogens could use such motifs to inhibit host chromatin processes, or use these motifs to co-opt cellular machinery.

Support for this has emerged from studies involving inhibitors (I-BET and JQ1) of a class of chromatin reader proteins known as the BET proteins. The BET proteins are a family of bromodomain containing proteins that bind to acetylated histones (particularly to histone H4) (Filippakopoulos et al., 2012). The BET inhibitors function as structural 'mimics' of acetylated histone H4 tails and exert their effects by competing with histone

binding to the BET bromodomains (Filippakopoulos et al., 2010; Nicodeme et al., 2010). Treatment of cells with either I-BET or JQ1 was sufficient to block BET recruitment to and function on chromatin (Dawson et al., 2011; Filippakopoulos et al., 2010; Nicodeme et al., 2010). Indeed, inhibition of BET proteins by I-BET resulted in the repression of the transcriptional response to LPS stimulation in macrophages (Nicodeme et al., 2010).

Mimicry is difficult to counter.

Third, mimicry, in itself, is a particular successful survival strategy for pathogens against their host. Typically, pathogen-host interactions can be exemplified by a simple arms race where host adaptation to virus is counteracted by pathogen adaptation, and *vice versa* (Daugherty and Malik, 2012). In the case of mimicry however, the spectrum of potentially useful adaptations that would facilitate host escape from mimicry are limited by the need to maintain other host–host interactions and functions(Elde and Malik, 2009). In addition, host defenses against pathogenic mimicry are further confounded by the fact that the host must now be able to not only recognize the offending pathogen-derived molecule, but must similarly be able to tell it apart from itself (Elde and Malik, 2009).

This is especially so in the case of the histone proteins. Given the vast number of interactions the histone tail motifs coordinate on chromatin and the crosstalk in which they engage in, histone tail mutations are not well tolerated. Indeed, it was recently shown that expression of histone H3 bearing missense mutations (lysine 27 to methionine, K27M) was sufficient to cause the loss of tri-methylated H3K27 throughout the cell, even in the presence of wild-type histone H3 protein (Lewis et al., 2013). While these observations do not preclude the possibility of compensatory mutations occurring in

histones during a host-pathogen arms race, the overall probability of that occurring is not likely to be high.

1.10 Hypothesis

Based on the above observations, we propose that mimicry of histone-derived motifs is likely to be a viable strategy by which pathogens can dysregulate host cell chromatin state and gene expression during infection. Specifically, pathogen histone mimics could exert their effects on host gene expression either by inhibiting the binding cellular factors from chromatin. Alternatively, pathogens might use mimicry to gain access to and co-opt the host chromatin processes. Understanding how pathogens can exploit the histone code will give insight into cellular processes, and may also be instrumental in developing therapeutics.

CHAPTER 2: MATERIALS AND METHODS

2.1 Cells and viruses

A549 and Madin–Darby canine kidney (MDCK) cells were obtained from the American Type Culture Collection (ATCC) and were maintained in Dulbecco's minimal essential medium (DMEM) (Gibco, Invitrogen) supplemented with 2mM Glutamine (Gibco, Invitrogen), 10% Fetal Bovine Serum (Hyclone) and penicillin-streptomycin (Gibco, Invitrogen). The A/Wyoming/3/2003 (H3N2), A/Puerto Rico/8/1934 (H1N1) viruses and the Flag-NS1 strains were propagated on MDCK cells. The A/Puerto Rico/8/1934(ΔNS1) (PR8/ΔNS1) virus was propagated in NS1-expressing MDCK cells. Vesicular stomatitis virus (VSV) (Indiana strain) was propagated in BHK cells.

2.2 Virus infections

A549 cells were plated at 5 x 105 cells per well in a 6-well plate 16-18 hours prior to infection. For infections, A549 cell monolayers were washed once in PBS, before being inoculated with 200 ul of influenza virus diluted to the appropriate MOI in PBS/0.3% BSA. Virus was allowed to adsorb onto the monolayer for 1 hour at 37°C, and plates were rocked gently every 10-15 minutes to ensure that the cell monolayer did not dry out. After viral incubation, the remaining virus was aspirated, and 2 ml of fresh growth medium was added back to the wells. Cells were then collected at various times post-infection for further assays.

2.3 Virus Growth Curves and Plaque Assays

Confluent A549 cells were infected at a multiplicity of infection (MOI) of 0.01. Following infection, cells were maintained in DMEM containing 0.3% Bovine Serum Albumin (BSA) and 0.375 μg/ml of tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin (Sigma). At the indicated time post infection, cell culture supernatants were collected. Viral titers at each time point were then quantified by plaque assays on MDCK cells.

For quantification of virus, dilutions of viral stocks or culture supernatants of the infected cells were adsorbed for 1 hour at room temperature onto layers of confluent MDCK cells. The infected MDCK cells were then overlaid with a 2 ml solution of DMEM containing 0.3% BSA, 25mM HEPES buffer (Gibco, Invitrogen), 2mM Glutamine (Gibco, Invitrogen), penicillin-streptomycin (Gibco, Invitrogen), 1 μg/mL TPCK-trypsin and 1% agar (LP0028, Oxoid). Plates were then incubated 48 to 72 hours until plaques could be observed. Plaques were then fixed in a solution of 7% formaldehyde, before being visualized by crystal violet staining.

2.4 Generation of Flag-NS1 viruses

The NS1 and NEP open-reading frames (ORFs) on the Influenza NS segment share a common N-terminal sequence. As such, to attach the Flag tag specifically to NS1 without disrupting the NEP ORF, the NS segment was modified as follows. The first 90 nucleotides of the 3'vRNA, with all the ATG start codons deleted, served as the 3' vRNA packaging signal. The endogenous splice donor site for NEP was left unchanged. The 3'vRNA packaging signal was followed by the 3XFlag sequence

(MDYKDHDGDYKDHDIDYKDDDDK) and the NS1 ORF with stop codon. Two silent mutations in the endogenous splice acceptor site in the NS1 ORF (TTCCAGGACATA) were introduced to prevent splicing at this site (TTCCCGGGCATA) as described previously (Varble et al., 2010). The Flag-NS1 ORF was followed by a new splice acceptor site that corresponds to the 459-527 nucleotides of the wild type NS segment, and the entire NEP-ORF with ATG. In this design, the 3XFlag-NS1 and NEP are generated from the unspliced or spliced mRNAs, accordingly. The deletion of the hPAF1 binding sequence was generated by introducing a stop codon after amino-acid 220 of the NS1 coding sequence. The modified NS segments were generated using fusion PCR and cloned into a pDZ vector using SapI restriction sites(Quinlivan et al., 2005). Flag-NS1 viruses were generated using reverse genetics system (Fodor et al., 1999). The sequence of the NS segment in the Flag-NS1 viruses were confirmed by RT-PCR and sequencing. Titers of viral stocks were determined by plaque assay in MDCK cells. A schematic of the targeting strategy is shown in Figure 2.1.

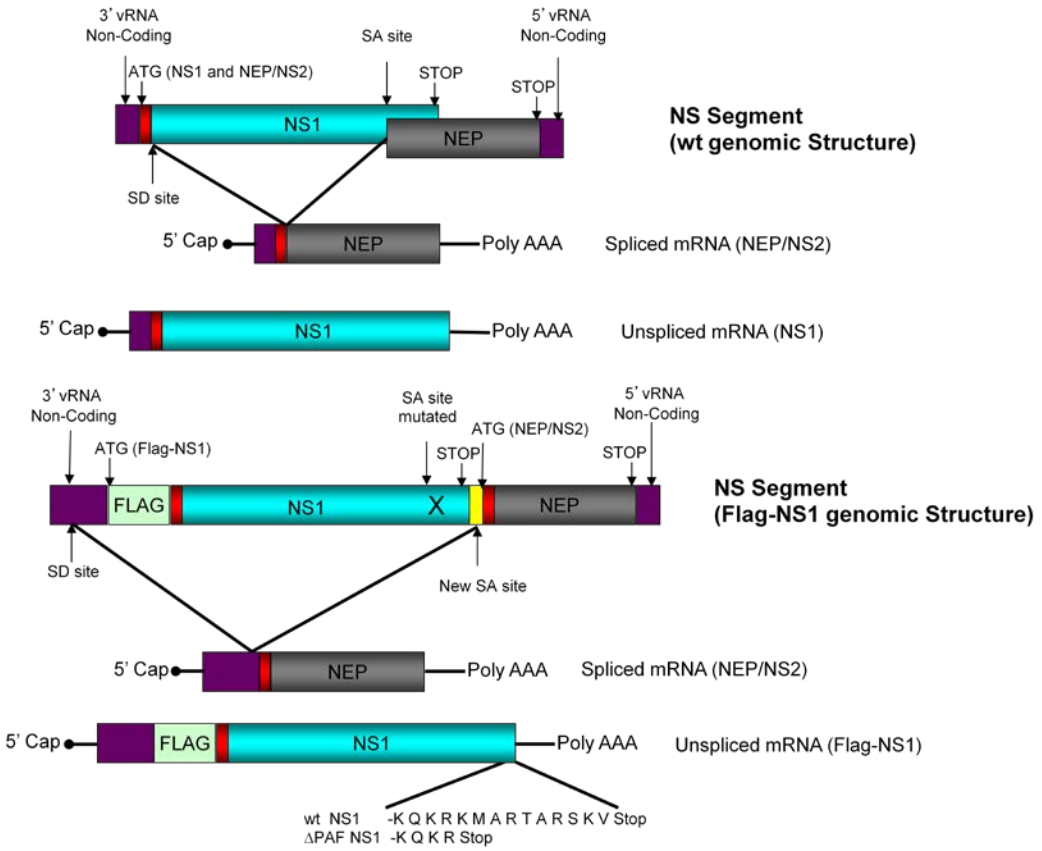


Figure 2.1: Generation of Flag-tagged NS1 virus Schematic of the genomic structure of the NS segment of wild-type Influenza A virus (wt) and the strategy used for the generation of the Flag-NS1 expressing influenza viruses. Mutation of the splice acceptor site of the NS segment enables selective Flag-tagging of NS1. SD, SA splice donor or acceptor sites, respectively.

2.5 siRNA mediated Knockdowns

Cells were transfected using Lipofectamine™ RNAiMAX Transfection Reagent (Invitrogen) according to the manufacturer's instructions. Cells were transfected with siRNA pools targeted to either human PAF1 (L-020349-01, Dharmacon), CHD1 (L-008529-00, Dharmacon) or a control non-targeting pool (D-001810-10-05, Dharmacon) at a final siRNA concentration of 50 nM. Transfected cells were used for further assays at

48 hours post transfection and gene knockdown efficiency was determined by quantitative PCR and/or Western blotting.

The sequences for the pooled human CHD1 siRNA oligonucleotides are as follows:

CACAAGAGCUGGAGGUCUA

GAUGAAGAUUGGCAAAUGU

CGAUCUCAUUUCUGAAUUA

GUACCGCUCUCCACUCUUA

The sequences for the pooled human PAF1 siRNA oligonucleotides are as follows:

GUGCCAUGGAUGCGAAAGA

GAGUACAACUGGAACGUGA

CUGUAGAAGAGACGUUGAA

CCACUGAGUUCAACCGUUA

2.6 Preparation of RNA-sequencing (RNA-Seq) libraries

RNA-Seq libraries were prepared with a protocol adapted from reference Rosenberg et al., 2011. Briefly, total RNA was extracted from infected A549 cells at different time points post infection using Trizol reagent (Invitrogen). Ribosomal RNA was depleted using the RiboMinus™ Eukaryote Kit for RNA-Seq (Invitrogen). Prior to fractionation, RNA was also treated with RNase-free DNase I (Qiagen) and purified using the RNeasy MinElute kit (Qiagen).

The RNA was fractionated in fragmentation buffer (40mM Tris acetate, pH8.2, 100mM potassium acetate and 30mM magnesium acetate) at 94 °C for 4.5 min. The fragmented RNA was reverse transcribed (Superscript III, Invitrogen) and then purified using the QIAGEN QIAquick PCR purification kit. The complementary DNA (cDNA) was then end-repaired using T4 DNA polymerase (NEB), DNA polymerase I, Large (Klenow) Fragment (NEB) and T4 PNK (NEB). End-repaired DNA was purified using the QIAGEN Qiaquick PCR purification kit. Klenow Fragment (NEB) was used to add 'A' bases to the 3'end of the DNA fragments before being purified by the QIAGEN MinElute PCR purification kit. Sequencing adaptor oligonucleotides (Illumina) were added with T4 DNA Ligase (NEB). Double-stranded cDNA libraries were then separated by electrophoresis through a 2% agarose gel, and fragments ranging from approximately 175 nt to 225 nt were excised and amplified by PCR with linker-specific primers (Illumina). The integrity and quality of RNA and cDNA were monitored throughout on the Agilent Bioanalyzer 2100. Ultra-high-throughput sequencing was performed on the Illumina Genome Analyzer II (GAII) by standard sequencing-by-synthesis reaction for 36-nt reads.

2.7 Immunofluorescence

A549 cells were cultured on coverslips overnight and then infected with the virus strains specified. At the indicated interval post-infection, cells were fixed in 3% paraformaldehyde (EMS) for 10 min at room temperature. Coverslips were washed in 1x PBS and blocked with blocking solution (1mg/ml BSA, 3% FBS, 0.1% Triton X100 and 1mM EDTA pH 8.0 in PBS) for 30 min at room temperature. Cells were then probed

with mouse monoclonal antibody against NS1 (diluted 1:300), or mouse monoclonal anti-Flag antibody (diluted 1:300) for 1 hr and detected by Alexa 594 conjugated Goat anti-mouse antibodies (Invitrogen). DNA was counterstained with DAPI.

2.8 Gene Expression Analysis by Microarray

Cells were infected with a virus strain that lacks NS1 (PR8/ΔNS1) at MOI 1 or stimulated with recombinant human IFN beta 1a (IFNβ1) (11415-1, PBL Interferon Source). Where cells were stimulated with IFNβ1, a concentration of 500 units/mL of cytokine was used. For infections with wild-type H1N1 influenza virus, the A/Puerto Rico/8/1934 (H1N1) strain was used at MOI 3. Infections with vesicular stomatitis virus (Indiana strain) were done at MOI 3. For Poly(I:C) stimulations, cells were transfected with Poly(I:C) at a final concentration of 2 μg/ml using the Lipofectamine2000 reagent (Invitrogen). Total RNA was isolated from infected, IFNβ1 stimulated or Poly(I:C) stimulated siRNA treated A549 cells using the RNeasy Kit (QIAGEN). 200ng of total RNA per sample was used to prepare biotin-labeled RNA using MessageAmp™ Premier RNA Amplification Kit (Applied Biosystems) and hybridized to HumanHT-12 v4 Expression BeadChips (Illumina). Data analysis was performed using the GeneSpring GX11.0 software (Agilent Technologies). Raw expression values were subjected to quantile normalization, and baseline transformation was performed to either the median of control samples for fold change analyses (see below), or to the median of all samples for comparisons between unstimulated siRNA treated cells.

To compare gene expression in siPAF- and control siRNA-treated cells, the normalized signal intensities of each microarray probe in the stimulated (infected or

IFNβ1 stimulated) samples was paired with and subject to baseline transformation against that of the corresponding un-stimulated sample that had been subject to the same siRNA treatment. An analysis of variance test (ANOVA) ($p < 0.001$), followed by a post hoc (TUKEY HSD) test and the indicated fold change cut offs were applied to identify probe-sets that showed statistically significant differences in expression upon stimulation for each siRNA treatment. Stimulation induced genes were defined as genes that are induced ≥2 fold ($p < 0.001$) in virus infected cells compared to un-stimulated cells in at least one siRNA treatment. hPAF1 dependent genes in virally infected and Poly(I:C) stimulated cells were defined as genes in which siPAF treatment induced a lower or greater (≥ 2 fold, $p < 0.001$) magnitude of response compared to siCtrl treated cells upon stimulation. hPAF1-dependent genes in IFNβ1-stimulated cells were defined as genes in which siPAF treatment induced a lower or greater (≥ 1.5 fold, $p < 0.001$) magnitude of response compared to siCtrl-treated cells upon stimulation. All p-value computations were subjected to multiple testing correction using the Benjamini Hochberg method.

For microarray analyses on the kinetic experiments with siPAF1 treated cells, the cells were infected with PR8/ΔNS1 at MOI 3. Total RNA was isolated used for microarray as described before. For the analysis, raw expression values were subjected to quantile normalization, and baseline transformation was performed to the median of all samples. Samples were compared via T-tests, since we did not include non-siRNA treated cells in these experiments. Entities that displayed > 2 fold change in expression ($p<0.01$) were subjected to further analyses.

Hierarchical clustering (Eisen et al., 1998) of data was performed and visualized using the Cluster and Treeview software (http://www.eisenlab.org/eisen/). Genes that are

represented by multiple probesets on the microarray are depicted by the average of those probesets in the heatmaps generated.

Functional analyses were conducted through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com). The Functional Analysis identified the biological functions that were most significant to gene lists generated from the microarray. Right-tailed Fisher's exact test was used to calculate a p-value determining the probability that each biological function assigned to that data set is due to chance alone.

2.9 Quantitative Real-Time PCR (qPCR)

Total RNA from stimulated cells was extracted using the RNeasy kit (QIAGEN) according to the manufacturer's instructions. RNA was DNase treated using the RNase free DNase kit (QIAGEN) and cDNA was synthesized using the First strand cDNA synthesis kit (Roche). qPCR was performed using SYBR green (Roche) or the LightCycler 480 Probes Master mix (Roche). The sequences of primers used are shown in Table 2.1.

Table 2.1: List of Primers used in this study. Listed are primers that have been used in this study. Hs: Homo sapiens; F: Forward primer; R: Reverse primer

Target	Primer name	Sequence	Use
NP mRNA (H3N2/Wyoming/2003)	H3N2_NP F H3N2_NP R	CCCAGGAAATGCTGAGATCG GTCGTACCCACTGGATACTG	mRNA detection
NS1 mRNA (H3N2/Wyoming/2003)	H3N2_NS1 F H3N2_NS1 R	TGGAAGGACCTCTTTGCATCA TCTTCGGTGAAAGCCCTTAGT	mRNA detection
PB2 mRNA (H1N1/PR/8/34)	PR8_PB2 F PR8_PB2 R	AGAGACGAACAGTCGATTGCCG ATCGCTGATTCGCCCTATTGAC	mRNA detection
NP mRNA (H1N1/PR/8/34)	PR8_NP F PR8_NP R	TATTGAGAGGGTCGGTTGCTCACA ACCAGTTGACTCTTGTGTGCTGGA	mRNA detection
M1 mRNA (H1N1/PR/8/34)	PR8_M1 F PR8_M1 R	GTGGCATTTGGCCTGGTA ATAGCCTTAGCTGTAGTGCTGG	mRNA detection
M1 mRNA (H1N1/PR/8/34)	PR8_M2 F PR8_M2 R	TAACCGAGGTCGAAACGCCTA GCCCTCCTTTCAGTCCGTATTT	mRNA detection
Universal primer (Influenza A)	Uni-12	AGCAAAAGCAGG	Reverse Transcription
Hs HPRT1	hs_HPRT1 F hs_HPRT1 R	TGAGGATTTGGAAAGGGTGT ACAGTCATAGGAATGGATCT	mRNA detection
Hs PAF1	hs_PAF1 F hs_PAF1 R	CAATTCCCACCGGACTCTG GTTTGTGCTGTTTCTCCAAGG	mRNA detection
Hs IFNB1	hs_IFNB1 F hs_IFNB1 R	CAGTCTGCACCTGAAAAGATATTATC GATTTCCACTCTGACTATGGTCCAGC	mRNA detection
Hs CCL5	hs_CCL5 F hs_CCL5 R	AAGCTCCTCTGAGGGGTTGA TTGCCAGGGCTCTGTGACCA	mRNA detection
Hs IFIT1	hs_IFIT1 F hs_IFIT1 R	TCCAGGGCTTCATTCATAT TTCGGAGAAAGGCATTAGA	mRNA detection
Hs IFIT2	hs_IFIT2 F hs_IFIT2 R	AGGCTTTGCATGTCTTGG GAGTCTTCATCTGCTTGTTGC	mRNA detection
Hs OAS1	hs_OAS1 F hs_OAS1 R	GATCTCAGAAATACCCCAGCCA AGCTACCTCGGAAGCACCTT	mRNA detection
Hs ISG15	hs_ISG15 F hs_ISG15 R	ACTCATCTTTGCCAGTACAGG CAGCTCTGACACCGACATG	mRNA detection
Hs MX1	hs_MX1 F hs_MX1 R	GTTTCCGAAGTGGACATCGCA GAAGGGCAAGTCCTGACACT	mRNA detection
Hs IFIT1 promoter/TSS	hs_IFIT1_prom F hs_IFIT1_prom R	CAGCTTACACCATTGGCTGCTGTT GGTTGCTGTAAATTAGGCAGCCGT	ChIP validation
Hs IFIT1 TES	hs_IFIT1_TES F hs_IFIT2_TES R	TTGGCTGACTTCACCTAGCTCACT CAAAGGACATAGAGGCACCCTGT	ChIP validation
Hs IFIT2 promoter/TSS	hs_IFI6_prom F hs_IFI6_prom R	TGATGCCCACACTTCATAGCTCCT TTTACTCGCTGCTGTGCCCATC	ChIP validation
Hs IFIT2 TES	hs_IFI6_TES F hs_IFI6_TES R	TGCTTGGGTTGTCTTCTCCTTCCT AAGAAGCGCTAGTGATCACCCTCA	ChIP validation

Commercially Available Primers

Targets	Company	Catalog Number	Format
Hs HPRT1	Applied Biosystems	4333768F	Taqman Probe
Hs ISG15	Applied Biosystems	Hs00192713_m1	Taqman Probe
Hs IFNB1	Applied Biosystems	Hs02621180_s1	Taqman Probe
Hs DDX58	Applied Biosystems	Hs00204833_m1	Taqman Probe
Hs CHD1	Applied Biosystems	Hs00154405_m1	Taqman Probe
Hs PAF1	Invitrogen	4331182	Taqman Probe

2.10 *In vitro* methylation assay

Methylation assays were performed as previously described (Nishioka et al., 2002)with minor variations. In brief, 300 ng of protein or peptide substrate and 100 ng of histone methyltransferases (HMT) were incubated with [^{3}H] SAM in HMT buffer (50 mM Tris-HCl (pH 8.5), 5 mM $MgCl_2$, 2 mM DTT) for 30 minutes at 37□C. The reaction was then immunoprecipitated for 1 hour with avidin beads (used with peptide substrates) or GST beads (used with GST-tagged substrates) and then washed extensively in BC150. This step minimizes non-specific radioactive incorporation. Eluted material was then subjected to PAGE, gel drying and exposed for radioactive signal detection. Set7/9 was a gift from Dr. Marc-Werner Dobenecker and purified SET1C was a gift from Dr. Tang Zhanyun.

2.11 *In vitro* acetylation assay

HAT reactions were performed in HAT assay buffer (50 mM Tris at pH 8.0, 10% glycerol, 50 mM KCl, 0.1 mM EDTA, 10 mM butyric acid, 1 mM dithiothreitol [DTT], 1 mM phenylmethylsulfonyl fluoride [PMSF]). Protein or peptide substrates (100ng) were incubated with [^{3}H]acetyl coenzyme A (CoA) and purified TIP60 (Gift from Dr. Xiao-Jian Sun) for 1h and affinity purified using analogous immunoprecipitation of the substrate before PAGE and detection on autoradiography film (see *in vitro* methylation assay).

2.12 Immunoprecipitation

Nuclear extracts from untreated and infected cells (pretreated with HDAC inhibitors when required) were denatured in Laemmli buffer (63mM Tris HCl, 10% Glycerol, 2% SDS, pH6.8) at 95□C for 10 minutes (with cycles of vortexing). The extract was then

diluted to a final concentration of 0.2%SDS in BC150 and sonicated with a Bioruptor (Diagenode). Proteins were then immunoprecipitated with Flag M2 antibody (Sigma)-coupled magnetic beads for 2 hours at 4□C. After extensive washing in BC300 and BC150 (last wash), the material was eluted with Flag-competing peptide at 37□C for 15 minutes (3 cycles) and the eluted material was combined and acetone-precipitated. Western blotting for NS1 modifications was followed by stripping and re-probing for loading control.

2.13 Peptide pull-down assays

Pull-down assays with extracts and recombinant proteins were performed as described previously(Wysocka, 2006). Nuclear extracts were prepared from HEK293 cells using the Dignam protocol (Dignam et al., 1983). 10^8 cells were used per pull-down assay. Salt and Triton-X100 concentrations were 250mM and 0.2% (v/v), respectively. Fractions from nuclear extracts fractionated on Heparin column, were pre-cleared with avidin beads and then incubated with biotinylated-peptide pre-bound to avidin beads for 3h at 4°C. Approximately 2µg of peptide was used per pull down. Beads were washed eight times with BC300 containing protease inhibitor cocktail (Roche). Bound proteins were eluted from the resin using 100mM glycine, pH2.8 and run on Micro-Spin Columns (Pierce, 89879). Eluates were combined, neutralized, and analyzed by SDS–PAGE. A similar procedure was used for peptide pull down using purified protein (100ng) or reconstituted complex (1µg). All peptides were synthesized by the Rockefeller University Proteomics Resource Center.

2.14 Antibodies

Anti-dimethyl NS1 antibody (NS1me2) was raised in rabbits against peptides (220-230) bearing pre-methylated K229 residue. Methyl specific antibodies were purified first by pre-absorbing serum (1 out of 8 rabbits showed highly reactive methyl-specific serum at the second bleed after peptide injection) to a matrix containing unmodified peptides, followed by purification on a NS1me2 column. Anti-PB1 and anti-NP are custom made antibodies kindly provided by P. Palese. Mouse anti-Flag is from Sigma (A8592); antibody against hPAF1C subunits were all purchased from Bethyl laboratories: PAF1 (A301-047A); CTR9 (A301-385A); LEO1 (A310-048A); RTF1 (A300 179A); Parafibromin/CDC73 (A300-170A); as well as CHD1(A301-218) and SMARCAL1(A301-086). Anti-H3K4me3 was purchased from Millipore (17-614), while Anti-RNA Pol II CTD (Ab5408) was purchased from Abcam. GST antibody was from Roche (RPN1236V).

2.15 Differential salt extraction

A549 cells were seeded and subsequently infected with influenza virus. Cells were then collected and nuclear pellets were prepared. Nuclear proteins were extracted from these pellets by using increasing concentrations of NaCl from 10 mM up to 2 M in BC buffer. Eluted materials were resolved on PAGE and immune-blotted with the specific antibody.

2.16 Chromatin-immunoprecipitation

We used a slightly modified version of described protocols to perform crosslinking ChIP (Barski et al., 2007; Lee et al., 2006). Approximately 10 million cells were used for each ChIP (5 times more material was used for Flag assay due to reduced performance of anti-Flag antibody on cross-linked material). In brief, uninfected or influenza infected A549 cells were fixed with 1% formaldehyde. The cross-linking reaction was stopped after 10 min by the addition of 2.5M glycine to a final concentration of 0.125M in the reaction. Cells were collected via scraping in ice cold PBS supplemented with protease inhibitor cocktail (Sigma). Cross-linked chromatin was subjected to sonication with the Biorupter (Diagenode), where we optimized sonication conditions to generate DNA fragments approximately 300-500 bp in length. Immunoprecipitations were carried out using antibodies pre-bound to either Invitrogen Dynal magnetic beads (Invitrogen Dynabeads anti-mouse M-280 #112-02, or Dynabeads anti-rabbit M-280 #112-04, or Dynabeads Protein A #100-02D). Following an overnight incubation, chromatin bound beads were washed 8 times in a modified RIPA wash buffer (50 mM HEPES-KOH pH 7.6, 100 mM LiCl, 1 mM EDTA pH 8.0, 1% NP-40, 0.7% Na-Deoxycholate) before being eluted into TE buffe containing 1% SDS 50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS). Protein bound chromatin complexes were then subject to overnight cross-link reversal at 65°C. After RNase and proteinase K digestion, ChIP DNA and input DNA were purified using the QIAquick PCR purification kit. DNA was eluted in TE buffer before being subject to downstream analyses.

2.17 ChIP-Sequencing

To prepare ChIP-sequencing libraries, we used 30ul of ChIP DNA and repaired DNA ends to generate blunt-ended DNA using the Epicenter DNA ENDRepair kit (Epicenter Biotechnologies, cat# ER0720). End-repaired DNA was purified using the QIAquick PCR purification kit (28104). Following DNA end Repair, we added A bases to the 3′ end of the DNA fragments using Klenow Fragment (NEB M0212L), and purified DNA using the QIAGEN MinElute kit (28004). We ligated Illlumina/Solexa adapters (#FC-102-1003) to DNA fragments overnight, using T4 DNA ligase (NEB M0202L). Following overnight ligation, we purified adaptor-ligated DNA fragments with the QIAGEN MinElute kit. To generate the final libraries for sequencing, we performed 18 cycles of PCR with Illumina/Solexa primers 1.0 and 2.0. We checked for fragment size by loading 1/10 of our amplified library on a 1% agarose gel, and purified the remaining ChIP-seq library using the QIAGEN MinElute kit. Purified library DNA was used for cluster generation on Illumina/Solexa flow cells, and sequencing analysis was performed on an Illumina/Solexa Genome Analyzer II following manufacturer protocols.

2.18 GRO-sequencing

Transcriptionally active nuclei from infected or untreated A549 cells were prepared after swelling for 5 minutes the cells in ice-cold swelling buffer (10mM Tris (pH = 7.5), 2mM $MgCl_2$, 3mM $CaCl_2$). Pelleted cells were re-suspended in 1ml lysis buffer (10mM Tris (pH = 7.5), 2 mM $MgCl_2$, 3mM $CaCl_2$, 10% glycerol, 0.5% NP40, $2U/ml^{-1}$ SUPERaseIN (Ambion) and pipetted 20 times with a P1000 tip with the end cut off to reduce shearing. Volume was brought to 10 ml with lysis buffer and nuclei were

pelleted at 600*g* for 5min. Nuclei were washed in 10ml lysis buffer and re-pelletted. A small aliquot was taken for Trypan blue staining to check that lysis occurred and nuclei were still intact. Nuclei were resuspended in 1ml freezing buffer (50mM Tris-Cl (pH = 8.3), 40% glycerol, 5mM $MgCl_2$, 0.1mM EDTA) using a P1000 tip with the end cut off and re-pelletted and re-resuspended in 500µl of freezing buffer and aliquoted into 100µl aliquots and frozen in liquid nitrogen. GRO-Seq libraries were then prepared as described previously (Core et al., 2008).

SSPE, NaCl, KCL, EDTA, and water are DEPC treated, while SDS, Sarkosyl, DTT, Tween, Tris buffers, PVP, NaOH were made with DEPC treated water, then filter-sterilized. Buffers used for immunoprecipitation contain superRNAsIN (1µl per 5 ml buffer) (Ambion) to block degradation that can occur during the experimental procedure.

2.19 Transcription Assay

In vitro transcription assays were done as previously described (Kim et al., 2010). In brief, we used an highly purified transcription factors (Pol II, TFIID, TFIIA, TFIIB, TFIIE, TFIIF, TFIIH, PC4, and Mediator) and a pML array template that contains p53-binding sites nearby the core promoter and generates 390-nucleotide transcription products. This system previously was shown to effect activator-dependent transcription (Kim et al., 2010). Purified proteins used for this assay were expressed recombinantly in a Baculovirus system in SF9 cells. Baculoviruses were generated according to the manufacturer's instruction (GIBCO-Invitrogen). To get purified proteins and/or complexes, SF9 cells were infected with combinations of baculoviruses. Proteins/complexes were then affinity purified on M2 agarose.

2.20 Bioinformatics used with high-throughput sequencing assays

Samples were sequenced in accordance with manufacturer protocols on GAIIx and HiSeq2000 instruments. Image data was analyzed in real-time by the onboard RTA software package.

Raw Data Analysis

Bcl files produced by RTA were converted to qseq files by Illumina's OLB software package, and qseq files converted to fastq for subsequent analysis.

ChIP-Seq Alignments

Sequencing reads were aligned to the Human March 2006 (NCBI36/hg18) assembly using the short-read aligner Bowtie (Langmead et al., 2009). Reads were aligned at 36bp allowing for 2 mismatches to the reference, reporting unique alignment locations only. RefSeq annotation data was downloaded from the UCSC table browser.

RNA-Seq Analysis

Sequencing reads were processed using Tophat (Trapnell et al., 2009), a junction mapping alignment program designed to identify splice junctions from RNA-Seq reads. Briefly, the program aligns reads to a reference genome, identifying regions of coverage that correspond to transcribed RNA. The underlying sequence of adjacent regions is joined together to create a spliced reference, and reads that did not initially align to the reference genome are aligned to identify sequencing reads that originated from potential splice junctions (e.g. exon-exon junctions).The Cufflinks (Trapnell et al., 2010) software package was used to perform gene expression level calculations and comparisons between RNA-Seq libraries prepared from uninfected cells and infected cells.

GRO-Seq Analysis

All sequencing reads were 51bp long. Reads that passed the internal Illumina quality filter were processed for adapter trimming, and reads which were longer than 15bp after adapter removal were retained. This subset of reads was aligned allowing 2 mismatches to the reference. Reads which were not trimmed of adapter sequence were truncated by 6bp at the 3'end to allow for potential incomplete adapter trimming, and aligned allowing 3 mismatches to the reference. Duplicate alignment positions were condensed to a single alignment entry to account for potential amplification biases. The adapter-trimmed and no-adapter alignments were merged for all subsequent analyses. All reads were aligned using the short-read aligner bowtie to the Human March 2006 (NCBI36/hg18) assembly.

Strand-specific coverage files were generated to differentiate between sense and antisense transcripts and to facilitate proper assignment of enrichment information for gene profiling.

GRO-Seq FPKM values were obtained by calculating the number of reads in the transcriptional unit and reporting per kilobase of gene length per million mapped reads.

Integrated profiles (see the following Integrated ChIP-profile) were made reflecting 3kb upstream from the TSS and 3kb downstream from the TES and 300 internal windows.

1000 genes were selected randomly from a group of 16,806 genes that had similar gene body enrichment level ranges relative to genes in Table 2. The GRO-Seq data from WT 12H was used for this selection.

Integrated ChIP-Seq profiles

Genes were profiled 2.5kb upstream of the TSS, through the gene body and 2.5kb downstream of the transcriptional end site (TES). Read counts were calculated in 100bp windows up and downstream of the TSS and TES, and each gene was segmented in 300 internal windows. Plots were made using a 1kb moving average. Values are read-normalized and reflect the number of reads observed in each averaged window.

Genes were selected by requiring a log2 fold-change increase in expression of greater than 2, PolII-Total ChIP-Seq enrichment increase throughout the gene body (600bp downstream of the TSS to 3kb downstream of the TES) of greater than >1.4-fold at 12 hours post infection compared to uninfected cells, and an H3K4me3 peak in either uninfected cells or cells 12 hours post-infection (as determined by MACS (Zhang et al., 2008) using custom settings for H3K4me3) within 3kb of the TSS. Additionally, genes passing these criteria were filtered out if the TES of genes with higher than 2 FPKM within 10kb of their TSS in an effort to minimize the effect of high RNA Pol II at the TES of highly transcribed genes.

For NS1 ChIP-Seq in A549 infected siPAF1, siCHD1 and siCtrl conditions, enrichment data was calculated for all RefSeq genes 5kb +/- TSS in 50bp windows, and anti-Flag ChIP-Seq enrichment values from uninfected cells was calculated similarly and subtracted. Values are read-normalized and reflect the number of reads observed in each window.

Peak Calling

To identify regions of ChIP enrichment, we used a custom JAVA/python/R peak calling algorithm that is based on the well known SPP (Kharchenko et al., 2008) and MACS2 (Zhang et al., 2008) algorithms. We first maximized the difference between the input and IP signal over the transcription start sites (TSS) regions of the reference genome to minimize false positive rates. A time series analysis of multiple time points of the PAF1 and Pol II ChIP-seq libraries was then perform to achieve the final list of Pol II or PAF1 bound regions.

CHAPTER 3: SUPPRESSION OF THE ANTIVIRAL RESPONSE BY AN INFLUENZA HISTONE MIMIC

3.1 Identification of putative histone mimics

The histone tails are composed of numerous overlapping short motifs that different effector proteins can recognize and bind to. In order to address our hypothesis that pathogens might have evolved similar motifs to interfere with host cellular regulation or to gain selective advantages, we designed an *in silico* screen. We focused on viral pathogens since they are obligate parasites which are known to rely heavily on host machinery. We focused our efforts on a set of known human viruses (Table 3.1)

For the screen, we first compiled a series of short overlapping motifs between 4 to 10 amino acids long deriving from the N-terminal histone H3 tail (Figure 3.1). The design and sequence of these motifs were influenced by the two following criteria:

(1) Search motifs were centered on known post-translational modification sites on the histone H3 tail. We reasoned that these residues were most likely to be encompassed within any putative motifs due to the fact that such sites often serve as important docking and regulatory sites for effector proteins.

(2) Residue charge and propensity to be modified were factored in the design of the putative motifs. For example, we allowed conservative serine –threonine substitutions since several known histone kinases are able to modify both serine and threonine residues on other substrates.

Table 3.1: List of Viruses used for the *in silico* screen. Shown here are the viral families, the genome type and the replication compartments of the viruses we used in our in silico screen.

Genome Type	Family	Replication Compartment	Virus
dsDNA (RNA intermediate)	Hepadnaviridae	Nucleus	Hepatitis B virus
dsDNA	Adenoviridae	Nucleus	Human adenovirus 1,2,35,5,54,7, A,B,C,D,E,F
	Poxviridae	Cytoplasm	Vaccinia virus, Variola virus
	Herpesviridae	Nucleus	Human herpesvirus 1,2,3,4,5,6A,6B,7,8
	Papillomaviridae	Nucleus	Human papillomavirus (types 116,10,101,103,108,109,112,121,126,128,129,131,132,134,135,136,137,140,144,16,166,26,32,34,4,41,48,49,5,50,53,60,63,6b,7,88,9,90,92,96)
	Polyomaviridae	Nucleus	BK polyomavirus, JC polyomavirus, KI polyomavirus, Merkel cell polyomavirus, WU Polyomavirus
ssDNA	Parvoviridae	Nucleus	Human parvovirus 4, Human parvovirus B19, Adeno-associated virus (1,2,3,4,5,6,7,8), Human bocavirus (1,2,3,4)
ssRNA (neg. polarity)	Hepatitis delta virus	Nucleus	Hepatitis D virus
	Rhabdoviridae	Cytoplasm	Rabies virus
	Filoviridae	Cytoplasm	Ebola virus - Mayinga Zaire 1976
	Paramyxoviridae	Cytoplasm	Measles virus, Mumps virus, Hendra virus, Nipah virus, Human respiratory syncytial virus, Human metapneumovirus
	Bunyaviridae	Cytoplasm	Bunyamwera virus, Hantaan virus, Sin Nombre virus, Rift Valley fever virus
	Arenaviridae	Cytoplasm	Lymphocytic choriomeningitis virus, Lassa virus, Junin virus, Machupo virus, Guanarito virus, Tacaribe virus, Sabia virus, Lujo virus, Mopeia Lassa virus reassortant 29
	Orthomyxoviridae	Nucleus	Influenza A virus (A/Hong Kong/1073/99(H9N2)), Influenza A virus (A/Korea/426/1968(H2N2)), Influenza A virus (A/Korea/426/68(H2N2)), Influenza A virus (A/New York/392/2004(H3N2)), Influenza A virus (A/Puerto Rico/8/1934(H1N1)), Influenza B virus, Influenza C virus (C/Ann Arbor/1/50)
ssRNA (pos. polarity)	Coronaviridae	Cytoplasm	Human coronavirus (229E, HKU1, NL63, OC43), Human enteric coronavirus strain 4408, SARS coronavirus
	Flaviviridae	Cytoplasm	Dengue virus (1,2,3,4), Hepatitis C virus, Hepatitis C virus (genotypes 2,3,4,5,6), Tick-borne encephalitis virus, Yellow fever virus, Modoc virus, Japanese encephalitis virus, West Nile virus, St. Louis encephalitis virus, Alkhumra hemorrhagic fever virus, Langat virus, Powassan virus, Wesselsbron virus, Usutu virus, Murray Valley encephalitis virus, Omsk hemorrhagic fever virus
	Togaviridae	Cytoplasm	Semliki forest virus, Sindbis virus, Venezuelan equine encephalitis virus, Western equine encephalitis virus, Eastern equine encephalitis virus, Chikungunya virus, Rubella virus
	Picornaviridae	Cytoplasm	Human enterovirus (100,107,98), Human rhinovirus B14, Theilovirus, Foot-and-mouth disease virus (types A, Asia 1, C, O, SAT1, SAT2, SAT3), Hepatitis A virus
	Caliciviridae	Cytoplasm	Sapovirus C12, Sapovirus Hu/Dresden/pJG-Sap01/DE, , Sapovirus Mc10
	Hepeviridae	Cytoplasm	Hepatitis E virus
ssRNA (DNA intermediate)	Retroviridae	Nucleus	Human T-lymphotropic virus (1,2,4), Human immunodeficiency virus 1 (1,2)
dsRNA	Reoviridae	Cytoplasm	Mammalian orthoreovirus 3, Adult diarrheal rotavirus strain J19, Rotavirus A, Rotavirus C

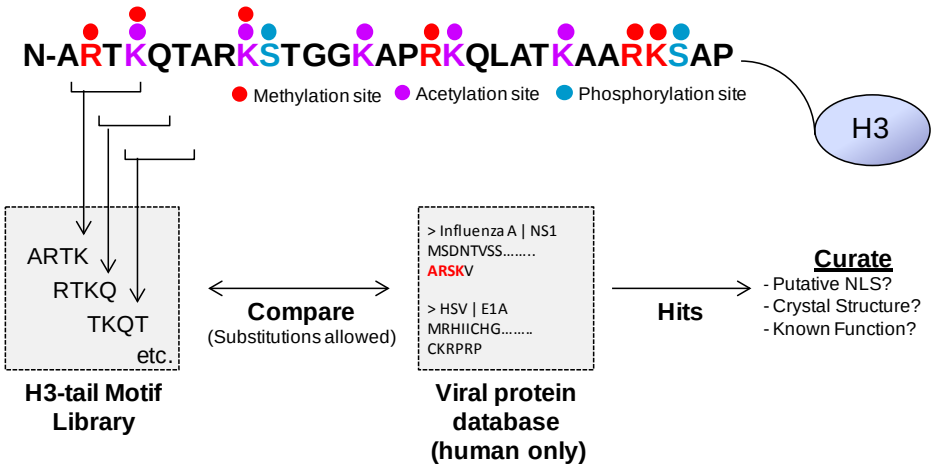


Figure 3.1: Strategy for *in silico* screen for putative histone mimics. A series of motifs deriving from the histone H3 N-terminal tail domain were compiled. This library of motifs was screened against a database of proteins derived from known human pathogens. We allowed certain conservative serine/threonine amino acid substitutions. Potential histone-mimics were scored for motif proximity to the protein terminus and manually curated for nuclear localization, structure and function in hosts. Histone diagram adapted from Zhang and Reinberg, 2001

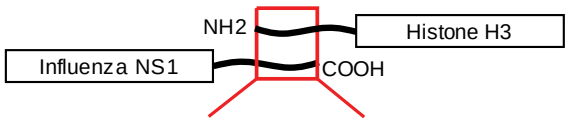
This compilation of motifs was then screened against a database of proteins derived from known human pathogens (Figure 3.1, Table 3.1) in order to identify potential viral histone mimics. Candidate proteins were ranked using a series of criteria. They were ranked by similarity of the identified motif to the original H3 sequence, as well as the proximity of the motif to either the carboxyl or amino terminus of the protein. In addition, the top hits for each motif were manually scored for known cellular localization as well as predicted or known function and structure. A list of some of our top candidate hits for H3K4-, H3K9- and H3K27-like motifs can be found in Table 3.2.

Table 3.2: Putative histone-like sequences found in known human pathogens. The table shows the top candidate viral proteins that bear histone H3K4-, H3K9- or H3K27-like motifs. The sequence of the identified motif and its distance from either the carboxyl (C-) or amino (N-) terminal is indicated.

Accession	VIRAL PROTEIN	MOTIF	Distance from terminal	Terminal	Localization
	H3K4- like sequences				
YP_308845.1	nonstructural protein 1 [Influenza A virus (A/New York/392/2004(H3N2))]	ARSK	4	C	Nuclear
NP_042931.1	DNA polymerase catalytic subunit [Human herpesvirus 6A]	ARSK	72	N	Nuclear
NP_819006.1	E2 protein [Semliki forest virus]	ARSK	31	C	Membrane
YP_012612.1	attachment glycoprotein G [Human metapneumovirus]	ARSK	23	N	Membrane/Cytoplasm
YP_006390078.1	truncated structural polyprotein [Semliki forest virus]	ARSK	106	C	Membrane/Cytoplasm
YP_001491557.1	NS5b [Hepatitis C virus genotype 3]	ARSK	97	N	Perinuclear
YP_081514.1	envelope glycoprotein B [Human herpesvirus 5]	ARSK	257	N	Membrane/Perinuclear
YP_001469632.1	HCV polyprotein [Hepatitis C virus genotype 4]	ARSK	494	C	Varied
YP_001469633.1	polyprotein [Hepatitis C virus genotype 5]	ARSK	494	C	Varied
YP_001469630.1	polyprotein [Hepatitis C virus genotype 2]	ARSK	494	C	Varied
	H3K9- / H3K27-like sequences				
AP_000576.1	pol [Human adenovirus 35]	ARKS	28	N	Nuclear
YP_002213842.1	DNA polymerase [Human adenovirus B]	ARKS	28	N	Nuclear
NP_040515.1	encapsidation protein IVa2 [Human adenovirus C]	ARKT	5	C	Nuclear
AP_000165.1	IVa2 [Human adenovirus 2]	ARKT	5	C	Nuclear
AP_000201.1	IVa2 [Human adenovirus 5]	ARKT	5	C	Nuclear
AP_000502.1	IVa2 [Human adenovirus 1]	ARKT	5	C	Nuclear
YP_001672011.1	E2 protein [Human papillomavirus type 88]	ARKS	52	N	Nuclear
YP_001974427.1	single-stranded DNA-binding protein [Human adenovirus D]	ARKT	72	N	Nuclear
YP_001129382.1	ORF29 [Human herpesvirus 8]	ARKT	126	N	Nuclear
YP_401712.1	BALF5 [Human herpesvirus 4]	ARKT	60	C	Nuclear

3.2 Influenza Non-Structural Protein 1 (NS1) from H3N2 subtype bears a histone mimic

One of our top candidate hits was the Non-Structural Protein 1 (NS1) sequence derived from the A/New York/392/2004(H3N2) strain of influenza A virus. Our screen showed that the NS1 protein carried the sequence 226-ARSK-229. This sequence strongly resembled the first 4 amino acids of the histone H3 protein 1-ARTK-4 (Figure 3.2). Strikingly, comparison of H3 and NS1 structural data showed that both these sequences were localized to the unstructured terminal tails of their respective proteins, indicating that the NS1 histone-like sequence was located in a similar structural context as the histone H3 tail.



Histone H3 1 ARTKQTARKS 10
NS1 221 KMARTARSKV 230

Strain	Year of appearance	NS1 COOH-terminal
H3N2	1968-2011	216 PKQKRKMART**ARSK**V 230
H1N1	1933-2011	216 PKQKRKMARTIRSEV 230
H5N1	2003	216 PNQKRKMARTIESEV 230
H1N1	1918*	216 PKQKRKMARTIKSEV 230
H1N1	2009*	216 PKQK---------- 219

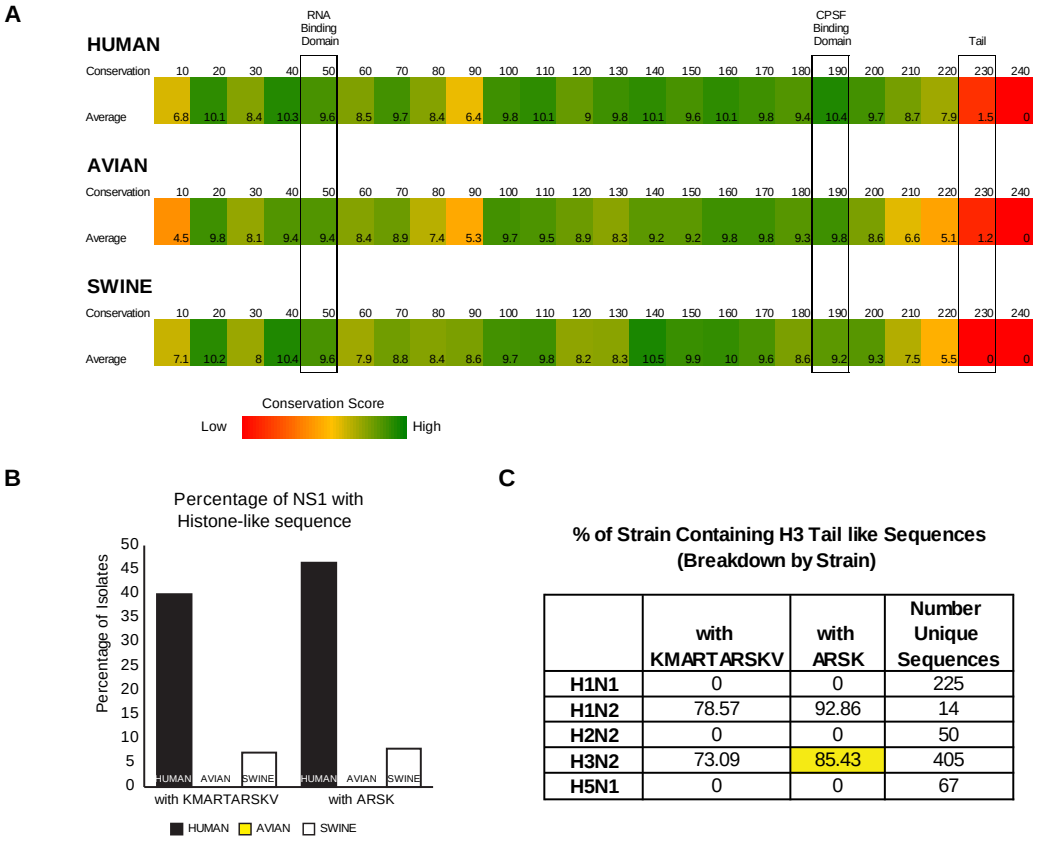
* Influenza pandemic strains

Figure 3.2: H3N2 Influenza A NS1 contains a histone mimic. The homologous carboxy-terminal NS1 and the amino-terminal histone H3 sequences are shown (red letters). The table displays C-terminal NS1 sequences of the influenza A subtypes.

We were also particularly intrigued by the NS1 protein it does not play a structural role in the influenza virion. Rather, its main function is to suppress the immune response during infection (Egorov et al., 1998; Garcia-Sastre et al., 1998; Kochs et al., 2007) and viruses lacking a functional NS1 protein are highly attenuated in immune-competent hosts (Garcia-Sastre et al., 1998). Depending on the Influenza virus subtype, it can vary from between 230 to 237 amino acids long (Palese P, 2007; Suarez and Perdue, 1998), although C-terminal truncations of 15-30 amino acids have also been reported (Suarez and Perdue, 1998). The protein can be divided into three functional domains: an N-terminal RNA binding domain (residues 1-73) (Chien et al., 1997; Hale et al., 2008b; Hatada and Fukuda, 1992), a C-terminal effector domain (residues 74-210) (Bornholdt and Prasad, 2006; Hale et al., 2008a). As alluded to earlier, the last 20 amino acids (hereafter referred to as NS1 'tail') of the effector domain (residues ~207-230) appear to be disordered, and are not observable in crystal structures of the NS1 effector domain (Hale et al., 2008a).

To understand whether the NS1 histone-like sequence was well conserved amongst influenza A virus isolates, we obtained 2753 unique full-length NS1 sequences from the NCBI influenza virus resource. Of these sequences, 1737 were derived from avian isolates, 250 were from swine isolates and 766 were from human isolates. Multiple sequence alignment of these proteins showed that the RNA binding, and C-terminal effector domain of the NS1 protein is generally very well conserved between viral isolates. In contrast, the C-terminal 'tail' domain, where the NS1 histone mimic was of the NS1 protein displayed poor sequence conservation (Figure 3.3A).

Influenza A viruses are typically categorized into subtypes based on hemagglutinin and neuraminidase serotype on their viral envelopes (Bouvier and Palese, 2008; Palese P, 2007). As such, when we broke our analyses down to subtype, we found that the NS1 histone-like sequence was highly conserved among influenza A isolates of subtype H3N2, but was not found in other subtypes (i.e. H1N1, H5N1) of human influenza A virus (Figure 3.2 and Figure 3.3C). This sequence was not found in any of the avian strains of influenza virus (Figure 3.3B). Altogether, this suggested that the NS1 histone-like sequence could be a strain- and human-specific virulence factor.



	with KMARTARSKV	with ARSK	Number Unique Sequences
H1N1	0	0	225
H1N2	78.57	92.86	14
H2N2	0	0	50
H3N2	73.09	85.43	405
H5N1	0	0	67

Figure 3.3: Conservation of NS1 histone-like sequence. A. Conservation of NS1 proteins within human, avian and swine isolates of influenza A virus. Green: High conservation score, Red: Low conservation score. **B.** Percentage of unique human, swine or avian NS1 sequences that carry the H3N2 NS1 tail ("KMARTARSKV") or the histone H3-like sequence ("ARSK") **C.** Abundance of NS1 histone-like sequence within different Influenza A subtypes.

3.3 NS1 is recognized by histone modifying enzymes

We hypothesized that the similarity of the NS1 tail to the histone H3 N-terminal tail would allow it to be recognized and modified by cellular histone modifying enzymes. To address this, we utilized three different approaches. In our first approach we incubated NS1 tail-peptides with known histone modifying enzymes and complexes. Notably, NS1

was methylated by the recombinant Set9 and Set1 complexes (Figure 3.4A), which specifically modify H3K4. Recombinant NS1 that carried a lysine to arginine mutation on residue 229 could not methylated by either Set1 or Set 7/9, showing that methylation occurred site-specifically. Similarly, acetyltransferase assays revealed that NS1 could also be acetylated at residue 229 by Tip60 complex, another known H3K4 acetyltransferase (Xhemalce and Kouzarides, 2010) (Figure 3.4A, top panel) In a similar vein, recombinant full-length NS1 expressed in bacteria incubated with the same cellular histone modifying enzymes and complexes was post-translationally modified *in vitro*. These modifications were also dependent on lysine 229 on the NS1 tail and mutation of lysine into an arginine resulted in the loss of post-translational modification (Figure 3.4A, middle panel).

We next sought to determine whether NS1 is modified in the context of a viral infection. To determine if methylation and acetylation of NS1 occurs during infection, cells were infected with either wild type (WT) virus or a virus bearing a lysine to arginine mutation at residue 229 in NS1 (K229R). NS1 methylation was detected with an NS1-methyl specific antibody that we raised (Figure 3.4B), whereas NS1 acetylation was detected by a pan-acetyl antibody. These experiments showed that NS1 is indeed methylated and acetylated in cells during viral infection. As expected, methylation and acetylation occurred on lysine 229 on the NS1 protein. NS1 isolated from K229R infected cells was not methylated or acetylated (Fig 3.4A, bottom panel). Together, these data show that NS1 is recognized by site-specific histone modifying enzymes, and can be post-translationally modified by these proteins.

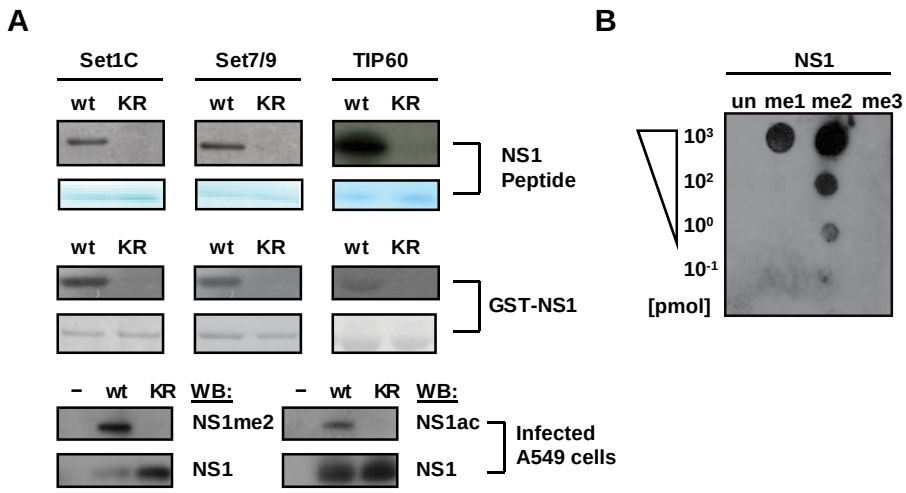


Figure 3.4: NS1 is modified by histone modifying enzymes. A. Methylation or acetylation of the NS1 peptide (top panel), the GST–NS1 protein (middle panel) or of viral NS1 in infected A549 cells (bottom panel) are shown. **B.** Characterization of the K229 methyl-specific anti-NS1 antibodies. Unmodified (un), or synthetically methylated (K229me1, me2 or me3) NS1 peptides were serially diluted at indicated concentrations and spotted on a nitrocellulose filter. The binding specificity was tested by dot-blot analysis using affinity purified methyl-specific rabbit NS1 antibody. The results show the specificity of the NS1K229 di-methyl specific antibody (anti-NS1me2). WT: wild-type NS1 sequence; KR: NS1 substrates where K229 is replaced by arginine; me1: Mono-methylation; me2: Di-methylation; me3: Tri-methylation; ac: Acetylation

3.4 NS1 is bound by PAF1 complex

Based on these results, we hypothesized that histone mimicry in NS1 would allow NS1 to interact with host histone-binding proteins, which in turn could be important for NS1 function. We thus sought to identify NS1 tail interactors by performing an unbiased peptide pull down screen. Briefly, nuclear extracts from HEK293 cells were fractionated over a heparin column and each resulting fraction was incubated with a peptide carrying the NS1 tail sequence. A scrambled control peptide, which had an identical amino acid composition to the NS1 tail, was used as a binding control. NS1 tail binding proteins

were separated by PAGE and visualized on a gel. Protein bands that were differentially bound between the NS1 tail peptides and its scrambled control peptide were than excised and subjected to tandem mass spectrometry (Figure 3.5A, left panel). A large number of peptides we found belonged to proteins that were derived from several complexes known to be associated with transcription and co-transcriptional activities. These included members of the human PAF1 transcription elongation complex (PAF1C), as well as the CHD1 chromatin-remodeling complex (Figure 3.5A, right panel).

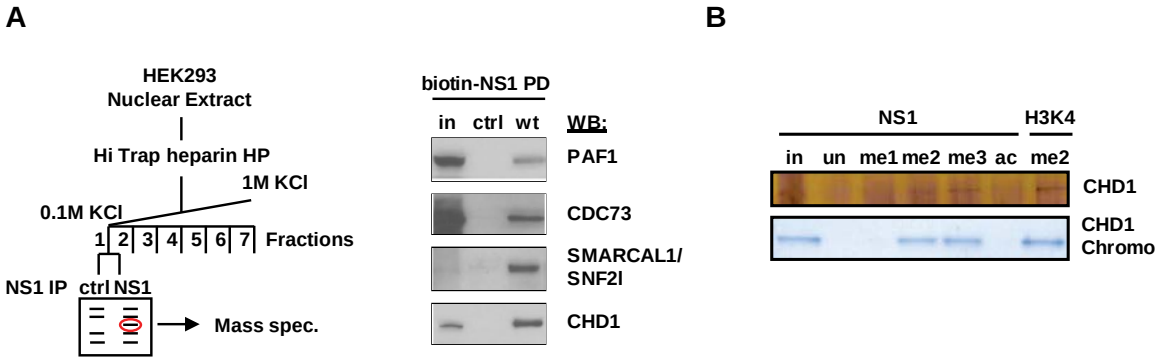


Figure 3.5: Identification of NS1 tail interacting proteins. A. Schematic depicting of affinity purification of NS1 "tail"-binding nuclear proteins (left panel). NS1 binding proteins were identified by affinity purification of HEK293 nuclear extracts and NS1 bound proteins were separated by PAGE and visualized by colloidal coomassie staining. Proteins that displayed differential binding to the NS1 "tail" and scrambled control peptides were extracted from the gel and analyzed by mass spectrometry. Association of the NS1 histone mimic with the PAF1C subunits and CHD1 in nuclear extracts (right panel). wt: Wild-type NS1 sequence; IP: Immunoprecipitation **B.** NS1 histone mimic binds to CHD1. Unmodified (un) or methylated (me) NS1 or H3(K4) peptides were incubated with recombinant full length CHD1 or the CHD1 double-chromodomain. Peptide binding was revealed by either silver staining or Coomassie staining.

We next used *in vitro* assays to reveal the primary binder of the NS1 protein. Of the proteins we identified, only CHD1 was previously shown to interact, albeit in a

methylation-dependent fashion to histone H3 (Sims et al., 2005). In support of the histone mimicry within the NS1 tail, CHD1 protein did not appear to directly interact with the unmodified NS1 tail peptide (Figure 3.5B). Instead, we saw that methylated forms of the NS1 tail peptide were able to interact well with human CHD1 protein, which had been identified as a reader of di- and tri- methylated H3K4 (Sims et al., 2005). Both full length human CHD1 and its purified chromodomain were found to interact with methylated NS1 tail peptide, but not to the unmodified or acetylated NS1 tail (Figure 3.5B).

We thus turned our attention to PAF1C because of several factors. The PAF1C has been shown to coordinate several steps in RNA polymerase II mediated transcription (Chen et al., 2009; Kim et al., 2010; Kim and Roeder, 2009; Mueller et al., 2004; Nordick et al., 2008), and has been shown to be a platform through which many other critical transcription regulatory factors may bind to (including CHD1) (Jaehning, 2010). PAF1C had also previously been implicated in the regulation of stress-induced genes in yeast (Betz et al., 2002; Kim and Levin, 2011), suggesting that it had a role in inducible gene expression. We were thus interested in further exploring the interactions between PAF1 and NS1

3.5 The PAF1 complex binds to the NS1 histone-like sequence

To prove that NS1 tail was directly binding to the PAF1 complex, we first assessed the ability of biotinylated NS1 tail peptides to bind to purified PAF1 complex in vitro. Our assay showed that the NS1 tail was sufficient to pull-down the PAF1 complex by immunoprecipitation (Figure 3.6B)

The human PAF1C is composed of six subunits – PAF1, CTR9, LEO1, RTF1, CDC73 and SKI8 – that bind together cooperatively to form the complex (Kim et al., 2010) (Figure 3.6A). However, both loss-of-function and mutational studies have suggested that there are distinct roles for the specific subunits within the complex, with each subunit binding to and interacting with different cellular partners(Betz et al., 2002; Mueller et al., 2004; Piro et al., 2012). Thus, to assess the primary binder(s) of the NS1 tail within the PAF1C, we incubated the biotinylated NS1 tail peptide with individual purified, Flag-tagged PAF1C subunits. The results of this experiment indicated that the PAF1 subunit is the primary binder of the NS1 tail. This binding was specific to the NS1 tail sequence, as a scrambled version of this peptide did not promote PAF1 binding. NS1-PAF1 association was also not strongly affected by methylation of K229 in the NS1 peptide, but was ablated by acetylation (Figure 3.6C).

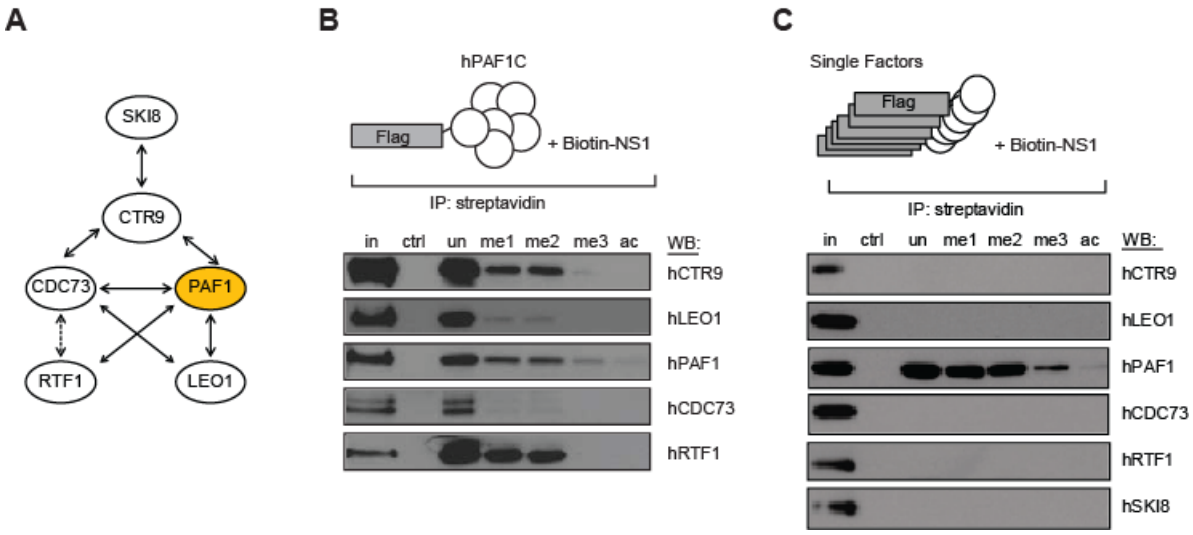


Figure 3.6: NS1 tail binding to recombinant hPAF1C and individual PAF1C subunits. A. Schematic of the human PAF1 complex (adapted from Kim et al., 2010) showing the binding interactions between the individual PAF1C subunits. Solid lines: stable associations; dotted lines: weak interactions **B.** Binding of NS1 peptides to recombinant hPAF1 complex. **C.** Binding of NS1 peptides to individual Flag-tagged hPAF1C subunits to NS1. For pull-downs, hPAF1 complex and individual hPAF1C subunits were prepared as described previously and incubated with the biotinylated NS1 peptides. Binding was assessed by western blotting. IP: immunoprecipitation; in: Input, ctrl: Control scrambled peptide; un: Non-modified; me1: mono-methylated; me2: di-methylated; me3: tri-methylated; ac: acetylated

Our experiments with the NS1 tail peptide also suggested that histone H3 might have similar interactions with PAF1C. Indeed, we found that unmodified and methylated H3 tail peptides, but not the scrambled control or acetylated peptides could bind to purified PAF1C or the PAF1 protein (Figure 3.7). Altogether these results highlight the similarity between the NS1 tail and the histone H3 tail.

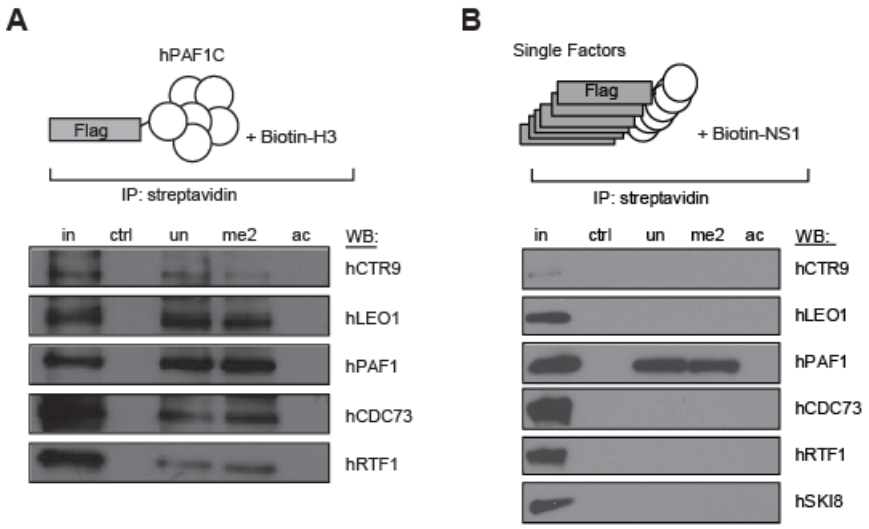


Figure 3.7: Histone H3 tail binding to recombinant PAF1C and its individual subunits. A. Binding of H3 peptides to recombinant hPAF1 complex. **B.** Binding of individual Flag-tagged hPAF1C subunits to H3. IP: immunoprecipitation; in: Input, ctrl: Contrl scrambled peptide; un: Non-modified; me1: mono-methylated; me2: di-methylated; me3: tri-methylated; ac: acetylated

3.6 NS1 is co-localizes with PAF1 on chromatin

In infected cells, NS1 is expressed with a predominantly nuclear localization (Figure 3.8A). It is expressed at high levels, with an estimated $5x10^5$ molecules per cell (Figure 3.8B). While it was possible that NS1 retains a nucleoplasmic role in the infected cell, the presence of the histone-like sequence in NS1, together with its ability to bind to complexes involved in transcription elongation, suggested that NS1 could also have chromatin related functions. Indeed, nuclear salt-extraction profiles of NS1 from infected cells revealed that NS1 is associated with chromatin (Figure 3.8B).

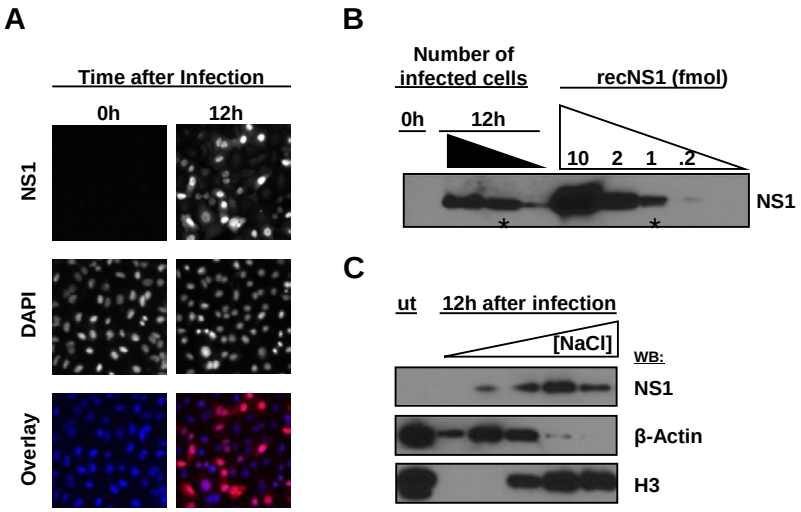


Figure 3.8: Nuclear localization and expression of NS1 protein in infected cells A. Immunostained NS1 (red) co-localizes with DAPI-positive (blue) nuclei in A549 cells at 12h after infection. **B.** The amounts of NS1 in serially diluted nuclear extracts of A549 cells were determined by Western blotting and compared to defined amounts of recombinant NS1. The amount of NS1 protein per cell was calculated based on Avogadro's equation ($NA = N/n$). C. The NaCl-elution profiles of NS1, β-actin and histone H3 are shown. The amount of indicated proteins in eluates was measured by Western blotting.

To facilitate further biochemical analysis of the NS1 protein, we knocked-in a Flag allele into NS1 using A/Wyoming/2003(H3N2) strain as a background (Flag-NS1) (Figure 2.1). This virus was infectious, and did not display any overt growth defects when compared to its untagged wild-type background, indicating that the 3X-Flag on the NS1 protein did not interfere significantly with viral biology (Figure 3.9).

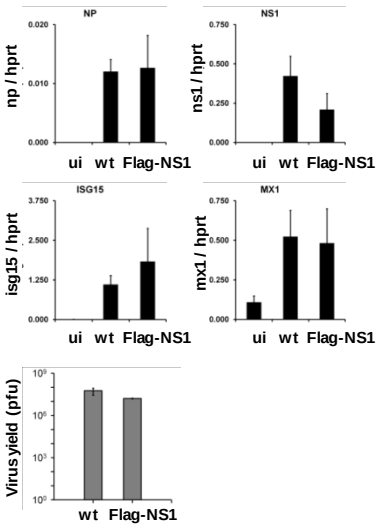


Figure 3.9: Flag-NS1 virus is infectious and displays no overt growth phenotype
Flag-NS1 recombinant virus supports infection and replication. A549 cells were infected with the wild-type or recombinant influenza virus that expresses Flag-tagged NS1. The virus functionality in single-cycle experiments was determined by expression of the viral nucleoprotein (NP) and NS1 in infected A549 cells (upper panel), by degree of up-regulation of the ISG15 and MX1 virus-induced genes in infected cells (middle panel) or yields of viruses propagated in MDCK cells (lower panel).

We infected cells with the Flag-NS1 virus and performed genome-wide chromatin immune-precipitation sequencing (ChIP-seq) experiments on Flag-NS1 binding. Since PAF1 has been implicated in transcription elongation, we reasoned that NS1 binding was likely to coincide with active gene transcription. We thus prepared ChIP-seq libraries for Pol II, PAF1 and H3K4me3 in the infected and non-infected cells. In addition, in order to identify genes that were induced upon infection, we performed RNA-sequencing. By cross-referencing the Pol II ChIP-seq libraries and the RNA-seq libraries, we were able to identify the subset of genes that were inducibly transcribed upon Influenza A infection.

Assessment of NS1 binding on infection inducible genes revealed that NS1 binding was enriched at the transcription start sites (TSS), with slight enrichments within the gene

body (Figure 3.10, top panel). This paralleled the enrichment profiles of both RNA polymerase II (Pol II) and PAF1 protein on these genes, indicating that NS1 was co-localized with those two factors (Figure 3.10).

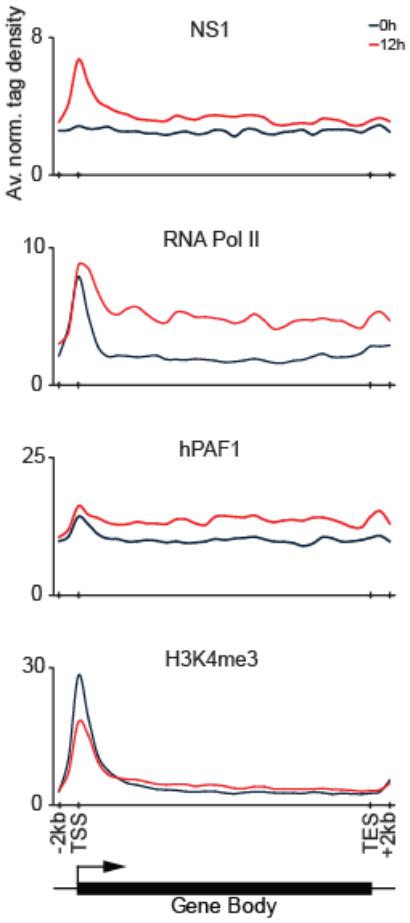


Figure 3.10: NS1 co-localizes with Pol II and PAF1 on chromatin The ChIP-seq profiles show the distribution of indicated proteins at inducible genes before (black line) and after (red line) infection. The induced genes were revealed by RNA-seq and ChIP-seq analysis of infected A549 cells. TSS and TES, the transcriptional start and end sites, respectively.

3.7 PAF1-binding activity of NS1 impacts hosts transcription elongation

The presence of NS1 at the antiviral genes suggested that NS1 would be in position to interfere with the recruitment and/or activity of PAF1 and Pol II at these loci. To differentiate between these possibilities, we performed ChIP experiments for Pol II and PAF1 in cells that were infected with either the wild type or ΔPAF viruses. Pol II and PAF1 abundance was then assayed at the transcriptional start sites (TSS) and transcriptional end sites (TES) of two known NS1-bound genes, IFIT1 and IFI6. In general, Pol II and PAF1 levels at these two genes were lower in wild-type infected cells, compared to ΔPAF-infected cells (Figure 3.11). However, the differences in Pol II and PAF1 binding between wild-type and ΔPAF virus infected cells were much more pronounced at the TES than at the TSS of these genes. Larger reductions of Pol II at the TES compared to the TSS are usually indicative of inhibition of transcription elongation, suggesting that this process could be impaired in wild-type infected cells.

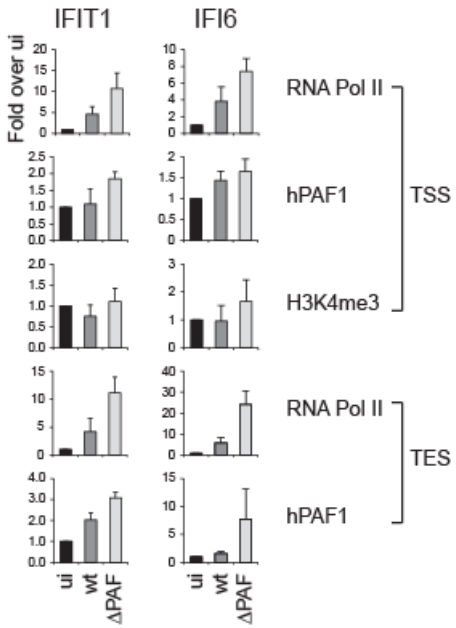


Figure 3.11: NS1 histone mimic is required for PAF1 and Pol II recruitment to chromatin PAF1, RNA Pol II and H3K4me3 levels at the TSS and TES of the induced genes in uninfected (ui) cells, cells infected with the wild-type (WT) or PAF1-binding mutant virus (ΔPAF). Data are representative of three independent experiments; error bars show the s.e.m.

ChIP experiments are useful for quantifying the amount of bound Pol II on the DNA, but they are unable to distinguish between transcriptionally engaged and inactive forms of Pol II. In fact, the decreased accumulation of Pol II at the at the IFIT1 and IFI6 TES as compared to the TSS, may either reflect inefficient transcription elongation, or be a result of an accumulation of transcriptionally paused and/or arrested Pol II at the TSS of the gene.

As such, to better understand the impact that the NS1 tail mimic had on antiviral gene transcription, we performed Global Run-on sequencing (GRO-seq) experiments on cells infected with either the wild-type or ΔPAF viruses. Metagene profiles of our GRO-seq data showed that in the absence of infection, the majority of the anti-viral genes are

transcriptionally silent, with low levels of active polymerases accumulating downstream of the TSS (Figure 3.12A). Many of these genes also retain a 5' peak of promoter proximal paused polymerase. Upon infection with either the wild-type virus or the ΔPAF virus, we saw an increase of active Pol II accumulate at both at the TSS and within the gene bodies of the antiviral genes. However, increase of active Pol II was significantly lower in wild-type infected cells as compared to that in ΔPAF infected cells (Figure 3.12A).

The differences in active Pol II accumulation between wild-type and ΔPAF infected cells also appeared to be specific to antiviral genes. Metagene profiles comparing wild-type and ΔPAF infected cells displayed no differences in Pol II accumulation at non-inducible genes, suggesting that the NS1-PAF1 interaction might be specific to infection-inducible genes only (Figure 3.12B).

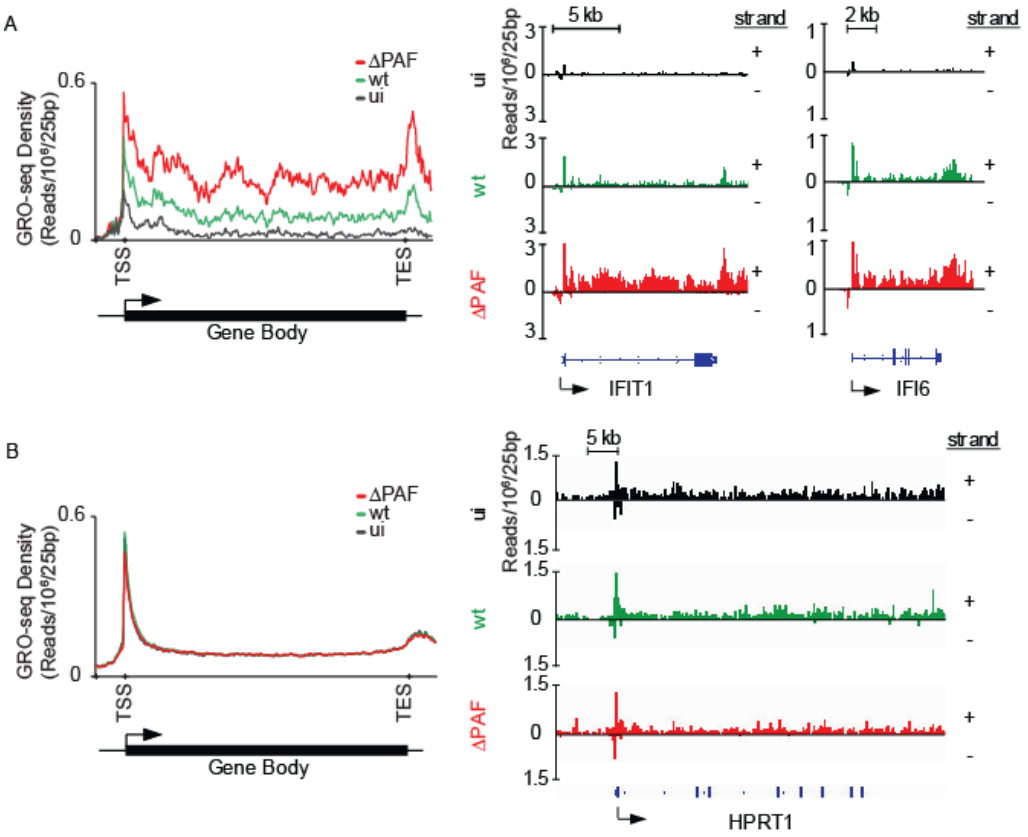


Figure 3.12: NS1 suppresses antiviral gene transcription in infected cells A. Left: the GRO-seq-measured RNA transcripts in uninfected (ui) A549 cells (black line) or cells infected with wild-type or ΔPAF virus (green and red lines, respectively) . Right: GRO-seq profile of IFIT1 and IFI6 genes in uninfected and infected cells. **B.** GRO-seq profile of A549-expressed genes that are not affected by virus infection (left panel) or of the HPRT1 gene (right panel). Reads from either DNA strands are indicated as +/− . The y axes display reads per million mapped reads per 25 bp.

3.8 NS1 histone like sequence affects antiviral gene expression

Our experiments with ChIP and GRO-seq suggested that transcription elongation might be impaired in cells that were infected with virus expressing wild-type NS1. To determine whether this defect was specific to the interaction between PAF1 and NS1, we performed *in vitro* transcription elongation assays where purified general transcription factors, transcription factors, co-activators and Mediator (Pol II, TFIID, TFIIA, TFIIB,

TFIIE, TFIIF, TFIIH, PC4 and Mediator) are used in conjunction with a chromatinized DNA template *in vitro* (Kim et al., 2010) (Figure 3.13A). In these assays, transcription elongation efficiency is measured by the generation of the full-length 390-nucleotide (nt) product from template.

Results from this assay indicated that NS1 strongly inhibited transcription elongation activity of PAF1C. As expected, addition of PAF1C to the reaction results in the accumulation of the 390 nt transcription elongation product (Figure 3.13B, lanes 1 and 2). This accumulation is reduced upon addition of purified NS1 to the reaction. In contrast, use of purified NS1 that lacks the PAF-binding (NS1(ΔPAF)sequence had little effect on PAF1c mediated transcription (Figure 3.13B, lanes 3). Altogether these results support the notion that NS1 tail sequence may function to interfere with PAF1C activity (Figure 3.13B, lanes 4).

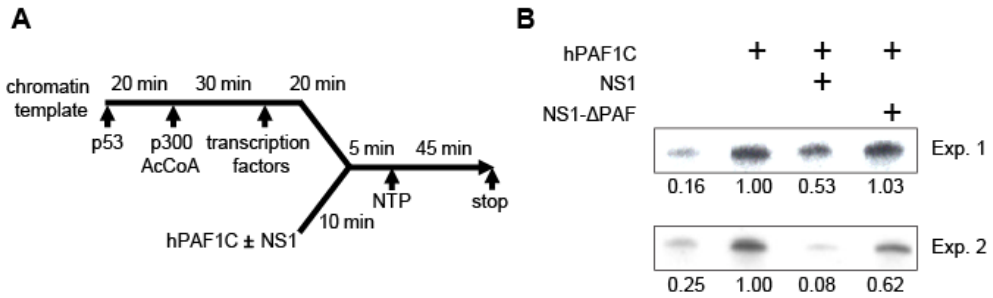


Figure 3.13: NS1 inhibits transcriptional elongation in vitro A. The full-length NS1 protein (NS1) or NS1 lacking the PAF1-binding sequence (NS1(ΔPAF)) was added to the RNA elongation reaction as indicated. **B.** The amount of the 390-nt RNA elongation product was quantified by ImageJ. The results of two independent experiments are shown.

3.9 PAF1 is required for the induction of the inflammatory response

The results in GRO-seq suggested that PAF1C might be involved in the up-regulation of the antiviral response. As such, we sought to study the impact that the PAF1C has on influenza-induced gene expression. To do so, we infected PAF1 deficient cells with NS1-null influenza virus (PR8/ΔNS1). We selected this virus as a precaution against potential NS1-PAF1 cross talk. This virus was also particularly useful because it induces an extremely strong antiviral response in host cells (Garcia-Sastre et al., 1998), allowing more subtle effects of regulatory factors on the antiviral response to be visualized easily (Shapira et al., 2009). For these sets of experiments, we chose to use a high multiplicity of infection (MOI) to reduce potential paracrine signaling events. We also limited our experiments to single-cycle infections.

Unexpectedly, despite being commonly thought of as a general transcriptional activity in the cell, PAF1 deficiency does not cause overt changes to gene expression in cells at steady state with expression levels of most housekeeping genes remaining unchanged between PAF1 siRNA (siPAF) treated and non-targeting control siRNA (siCtrl) treated cells (Figure 3.14).

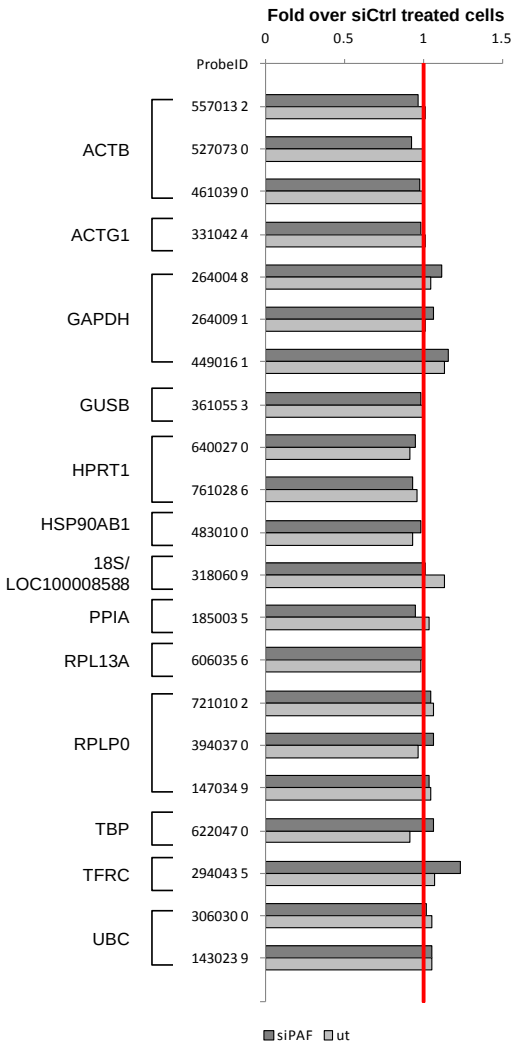


Figure 3.14: PAF1 is not essential for housekeeping gene expression. Expression levels of indicated housekeeping genes were determined by microarray analysis of RNA derived from un-transfected (ut), siPAF or siCtrl transfected cells. Results for individual probesets are shown for genes that are represented by multiple probesets on the microarray.

Infection of siCtrl and siPAF treated cells with PR8/ΔNS1 infection revealed that there were similar numbers of up- and down-regulated genes in both the siCtrl and siPAF treated cells. However, a subset of infection up-regulated genes appeared to be sensitive to PAF1 depletion. These genes, which were highly enriched in key antiviral genes

(including IL28A, IL28B, IL29, DDX58/RIGI), were induced to lower levels in siPAF treated cells as compared to siCtrl treated cells (Figure 3.15A). Further qPCR analysis on several of these genes validated our findings in microarray (Figure 3.15B).

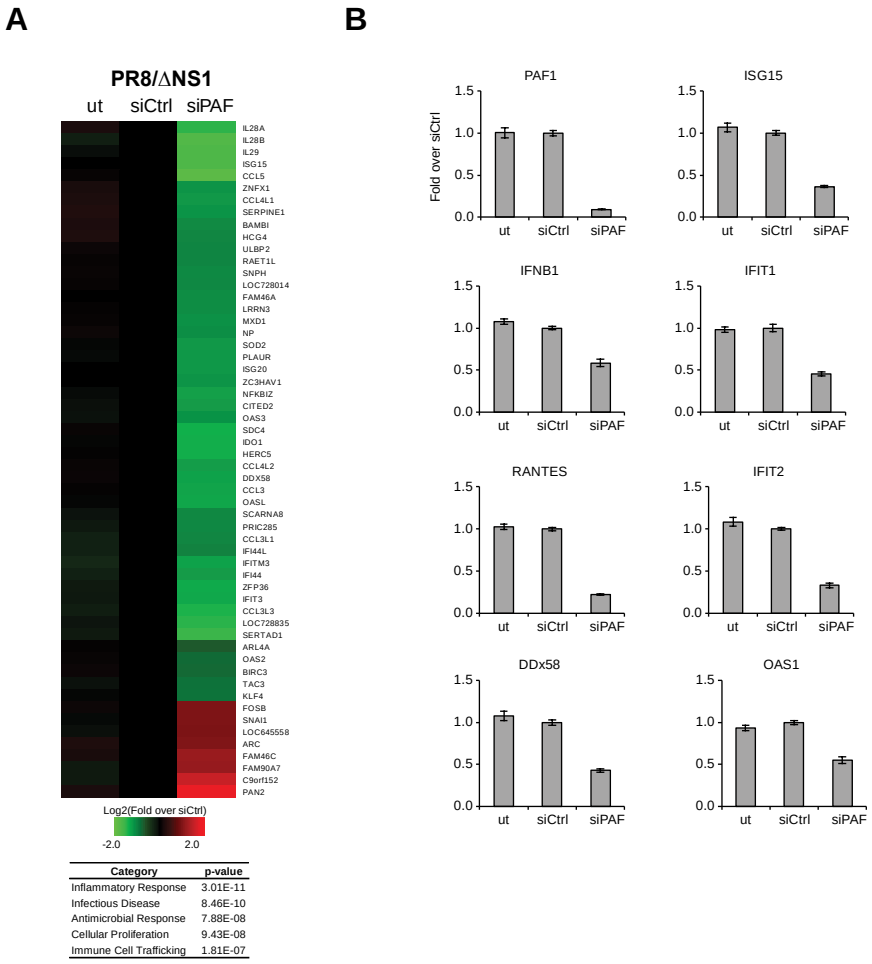


Figure 3.15: PAF1 controls antiviral response. A. The expression levels of mRNAs in influenza infected control (siCtrl) or PAF1-deficient (siPAF) A549 cells. The table shows top siPAF-affected gene categories as identified by IPA. ut, untreated with siRNA. **B.** qPCR analyses of anti-viral genes in siCtrl, siPAF and ut cells. The fold difference between levels of hPAF1 mRNA (upper panel) and host mRNA up-regulation were measured by quantitative real time PCR of RNA isolated from infected A549 cells.

To determine whether the changes we saw in siPAF-sensitive genes could be attributed to alterations in viral gene expression, we quantified the levels of viral mRNA

and viral genomic RNA in siPAF and siCtrl cells. We found no significant changes in the expression of virally-derived RNAs, suggesting that the differences we saw in antiviral gene expression were not likely to be a result of an altered potential of the virus to induce the inflammatory response in siPAF and siCtrl cells (Figure 3.16).

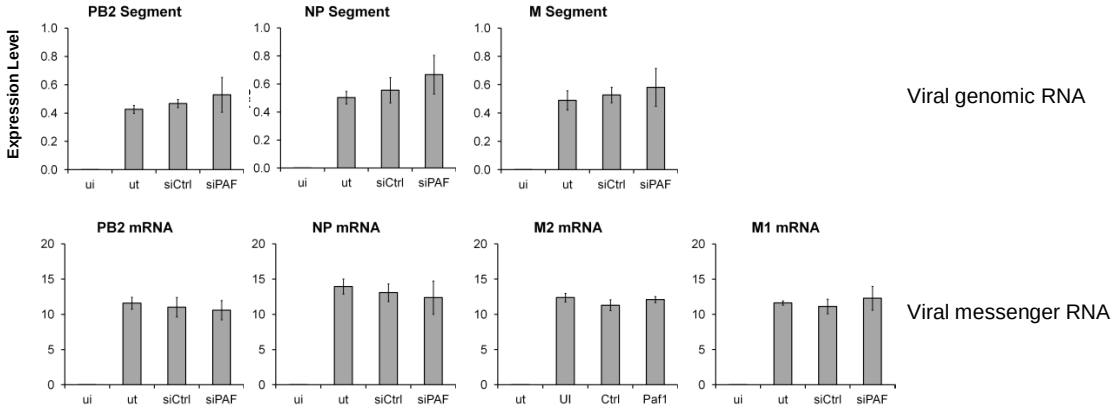


Figure 3.16: PAF1 does not control production of viral RNAs in PR8/ΔNS1 infected cells. The levels of the indicated influenza genomic (upper panel) or messenger RNAs (lower panel) were measured by qPCR analysis of RNA derived either from uninfected (ui), PAF1-deficient (siPAF1) or control (siCtrl) A549 cells infected with PR8/ΔNS1 viruses. Data are representative of 3 independent experiments. Error bars represent the S.E.M.

Similar to PR8/ΔNS1 infected cells, we found that antiviral gene expression was reduced during infection with the NS1 bearing parent strain of PR8/ΔNS1, A/PR/8/34 as well as during infection with a non-related virus vesicular stomatitis virus (VSV) (Figure 3.17). These results suggested that PAF1 requirement in the antiviral response was not specific to influenza virus. Thus, to understand if active viral replication was required for siPAF1 mediated effects on gene expression, we treated PAF1 deficient cells with defined amounts of either IFNβ (a major cytokine regulating the antiviral response) or Poly(I:C) (chemical mimetic of viral genomic RNA). (Figure 3.17) Analysis of these

experiments showed that similar to infected cells, the expression of IFNβ and Poly(I:C) induced genes was reduced in PAF1 deficient cells. Altogether, these data suggested that PAF1 was required for the proper induction of stimulus responsive genes.

Figure 3.17: hPAF1C controls antiviral gene expression in response to various stimuli. The expression levels of antiviral genes were measured by microarray analysis of RNA isolated from wild-type influenza H1N1 (left panel), vesicular stomatitis virus (VSV) (middle panel), IFNB1 treated (top right panel) or Poly(I:C) treated (lower right panel) A549 cells that were either not transfected (ut) or transfected with control (siCtrl) or hPAF1 (siPAF) specific siRNAs. The tables show the top five functional categories of the siPAF affected genes as identified by IPA.

The inability of siPAF1 treated cells to mount a full antiviral response during infection suggested that siPAF1 deficient cell populations would be highly susceptible to viral infection. Accordingly, when siPAF1 and siCtrl treated cells were infected at an m.o.i of 0.01 with PR8/ΔNS1 in a multi-cycle infection; we saw a log fold increase of viral yield (Figure 3.18). On the other hand, depletion of CHD1, which we had also found in our screen, had little effect on viral growth (Figure 3.18).

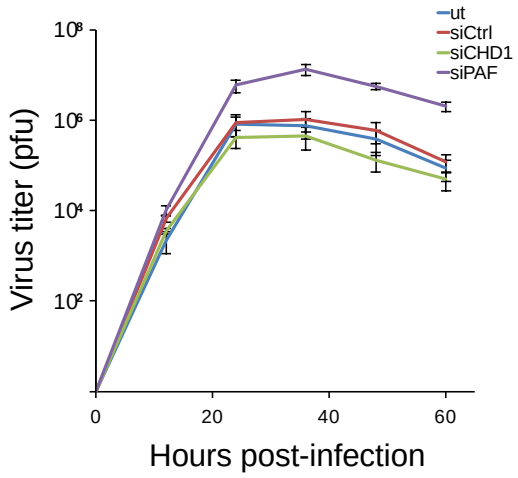


Figure 3.18: Dynamics of virus replication in control or PAF1-deficient A549 cells. Shown are the viral growth kinetics of Influenza infected cells. p.f.u. plaque-forming units. Data are representative of three independent experiments. Error bars show the s.e.m.

3.10 Conclusions

In summary, we found that the NS1 protein of the H3N2 subtype of Influenza virus carries a histone H3-like sequence in its C-terminal tail domain. This sequence is required for the virus to interact with the host PAF1C, and in doing so limit the host anti-viral response by impairing PAF1C function (Figure 3.19). Lastly, we have used this host-

pathogen interaction to uncover the central role of transcription elongator complex PAF1 in the proper induction of the immune response.

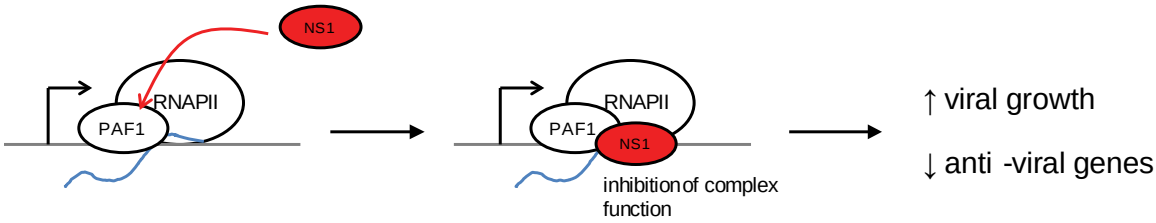


Figure 3.19: Putative Model of NS1: PAF1C interaction. NS1 histone-like sequence is required for NS1 to interact with the PAF1C. PAFC1 function is inhibited upon NS1 binding, resulting in reduced production of anti-viral genes, and increased viral spread.

CHAPTER 4: DYNAMICS OF PAF1 RECRUITMENT ON CHROMATIN DURING THE ANTIVIRAL RESPONSE

4.1 Preamble

Our studies involving Influenza NS1 and its histone mimic highlighted the specific role of PAF during viral infection. In fact, even though the PAF1C is considered to be a general RNAPII associated elongation factor, our analysis showed that the depletion of PAF1 in cells only had a very selective impact on the expression on anti-viral genes. As such, the goal of this part of the study was the elucidate characteristics of these PAF1-target genes. For these studies, we chose to use infection with PR8/ΔNS1 as our experimental system. This system was specifically chosen because this virus is a very potent inducer of the antiviral response, and whose selective impact on the antiviral genes can be ascribed efficiently. We predicted that PAF1 recruitment and activity is differentially regulated on the induced anti-viral genes compared to non-induced, non-infection related genes.

4.2 Dynamics and Specificity of PAF1 binding during infection

In order to understand the specific requirements of PAF1 recruitment to chromatin during infection, we performed ChIP-sequencing analysis of PAF1 binding sites from uninfected and PR8/ΔNS1 infected cells. We chose to use both early (4 hours post infection) and a late (12h hours post infection) time points in this analysis to gain better understanding of the dynamics of PAF1 binding.

To determine the effects of infection on genome-wide PAF1 binding we took an unbiased approach and used a customized peak caller program (see Methods) to identify all PAF1 bound regions across the genome. We did encounter some limitations in the program for detecting very broad regions of binding on chromatin and were unable to completely avoid instances where a single large region of PAF1 binding was broken into several smaller regions. Despite this, our approach appeared to be robust, with 98% of genes that were bound by PAF1 being associated with a maximum of three PAF1 peaks per gene. Two examples of the input raw sequence reads and the regions that have been identified as PAF1 bound is shown in Figure 4.1A.

Overall, we were able to map at least 9550 loci across the genomes which are associated with PAF1 binding in infected or uninfected cells. A majority (86%) of these loci were associated with genic regions, with 46% of the called peaks localized to intra-genic sites (i.e. within the gene body); 5% localized to areas around the transcription termination sites and 32% localized to areas around the transcription start sites and promoter regions of genes (Figure 4.1B). In contrast, only about 17% of these PAF1 bound loci were found in inter-genic regions. The mostly genic association of PAF1 is consistent with the known association of PAF1 with RNA polymerase II, and its function as a key player in transcription elongation and other co-transcriptional processes.

Next, we investigated the overall regulation of PAF1 binding throughout the course of infection. PAF1 localization across the genome appeared to be fairly stable during infection, with only 2 to 4% of loci showing differential recruitment of the PAF1 protein during infection (Figure 4.1C). In fact, only a total of 198 loci at 4 hours post-infection, and 403 sites at 12 hours post-infection showed at least a 2-fold change in PAF1 binding.

Of these sites, the majority of the changes occurred in intragenic regions and was associated with an increased level of PAF1 binding at the loci. On the other hand, the numbers of sites that lost or gained PAF1 binding within the TES, TSS and intergenic regions were similar to each other (Figure 4.1C).

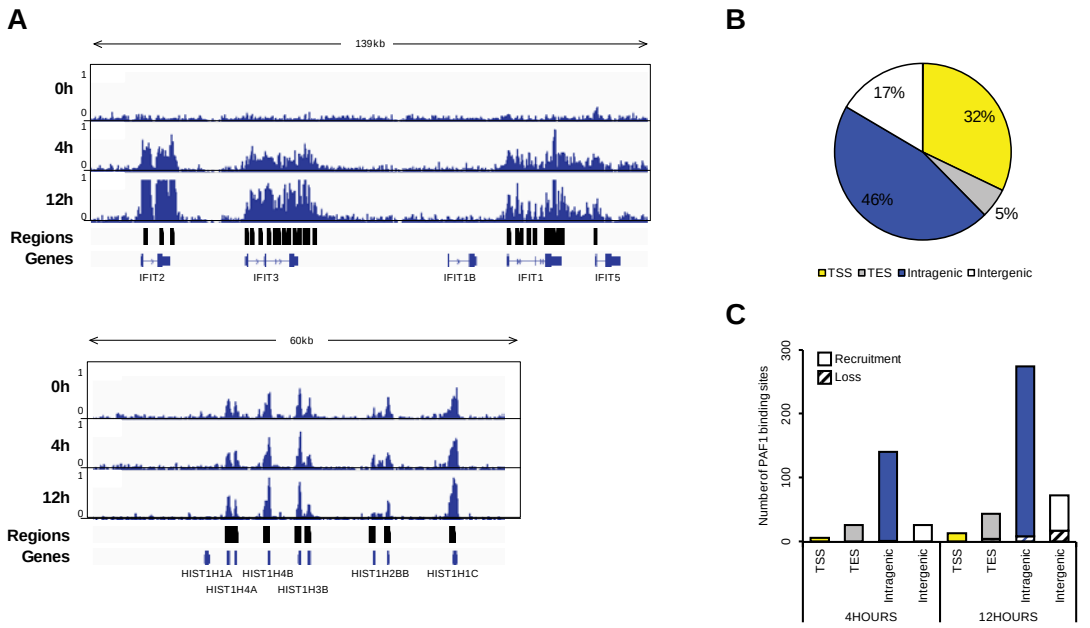


Figure 4.1: Genome-wide distribution of PAF1 binding. A. Genome browser view of PAF1 binding at the IFIT locus (top panel) and the HIST1H1C locus (bottom panel) in uninfected cells, and infected cells 4 hours (4h) and 12 hours (12h) post infection. PAF1 bound regions that were called using our customized peak finder program are highlighted in black boxes above the genes. **B.** Classification of PAF1 bound regions based on their overlap with transcription start sites (TSS), transcription termination sites (TES), intragenic regions (including introns and exons), as well as intergenic regions. **C.** Numbers of PAF1 bound loci that display at least a 2 fold increase in PAF1 binding at 4 hours or 12 hours post infection. PAF1 bound loci are classified as described in B.

4.3 PAF1 is strongly enriched on antiviral genes upon infection

We next scrutinized the PAF1 bound loci that displayed differential recruitment of PAF1 during infection. Since the majority of sites that show changes in PAF1 binding

during infection occur within genic regions (TSS, intragenic, TES), we decided to focus our efforts on gene-associated PAF1 bound regions.

At 4 hours post infection, genic loci that displayed at least a 2-fold change in PAF1 binding were associated with a total of 59 genes (Figure 4.2A). Of these 59 genes, 58 displayed an overall increase in PAF1 binding, whereas 1 gene displayed a loss in PAF1 binding. Similarly, a total of 126 genes were associated with a change in PAF1 binding at 12 hours post infection. Of these genes, 116 genes were associated with an increase in PAF1 binding, whereas 8 genes displayed a loss of PAF1 binding upon infection (Figure 4.2B, Left and Right panels).

To determine if genes that lose or gain PAF1 upon infection belong to identifiable groups, we carried out pathway analyses on these genes. Genes that displayed an increase in PAF1 binding either at 4 hours and/or a 12 hours post infection were highly enriched in genes associated with inflammatory and antimicrobial response (Figures 4.2C and 4.2D), which are both impacted by Type I interferon signaling(Stetson and Medzhitov, 2006) . Consistent with this, we also saw the presence of many known interferon stimulated genes (Schoggins et al., 2011) amongst the genes that actively altered PAF1 binding during infection (Figures 4.2A and B). In addition, we found an overrepresentation of NF-κB and IRF binding sites at the promoters of genes that had altered PAF1 binding during infection (Figure 4.2E). In contrast, these factors were not identified in a set of randomly chosen genes (Figure 4.2E). Since both IRFs and NF-κB are important for the activation of inflammatory genes, this observation further underscored the specific recruitment of PAF1 to anti-viral and inflammatory genes during infection.

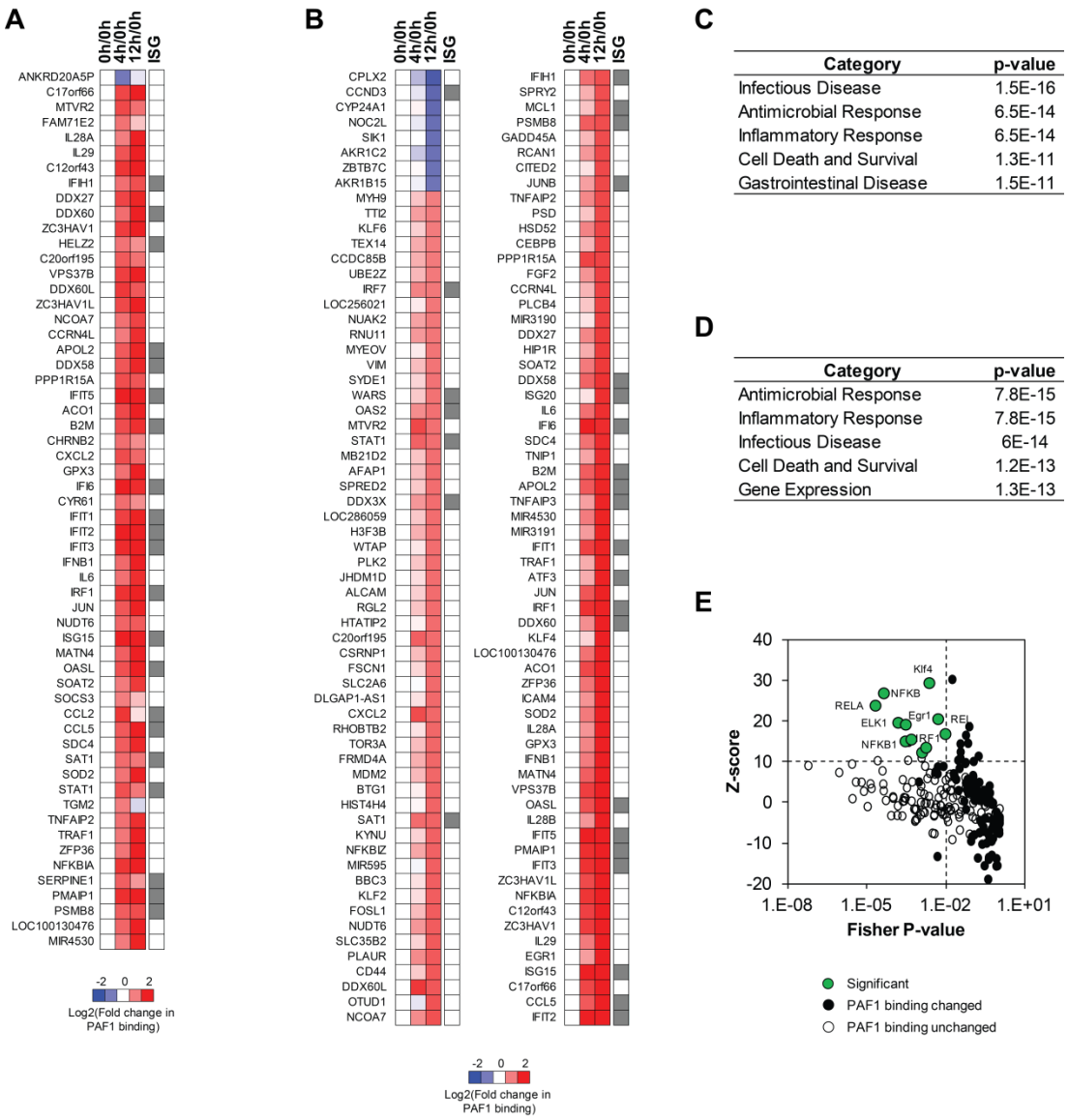


Figure 4.2: Changes in PAF1 recruitment at genes during infection. A. Genes that display ≥2 fold change in PAF1 binding at 4 hours post infection. Increased PAF1 binding is shown in red and decreased PAF1 binding is shown in blue. ISGs (Schoggins et al., 2011) are indicated by grey boxes on the right of the heatmap. **B.** Same as in A, but for PAF1 binding at 12 hours post infection. **C.** Top five functional categories associated with genes that are differentially bound by PAF1 (>2 fold) at 4 hours post infection as identified by IPA analysis. **D.** Same as in C, but for 12 hours post infection. **E.** Overrepresentation of transcription factor binding sites (TFBS) found -5kb to +5kb around the TSS. TFBS that received Z-score values >10 and p-values < 0.01 were considered to be significant, and are indicated by green circles. All TFBS associated with the 136 genes associated with differential PAF1 binding are displayed as black circles; All TFBS associated with a set of 136 randomly selected PAF1 binding genes are shown as white circles.

4.4 PAF1 recruitment is correlated with gene expression and Pol II recruitment

Our data on PAF1 binding indicate that the dynamics of PAF1 recruitment on chromatin during infection is highly specific and is strongly dominated by genes belonging to type I interferon signaling and hallmark genes of the antiviral response. Given the important role that the PAF1C plays as a transcription elongation factor, we hypothesized that PAF1 recruitment to these genes is critical for their expression. To address this, we utilized two independent approaches to confirm whether an increase in PAF1 binding at these genes was positively correlated with transcriptional activity of genes during infection.

The abundance of Pol II at a gene, particularly within its gene body, can be used as in indicator of transcriptional activity of a gene. We first wanted to determine if PAF1 binding was correlated with Pol II abundance across the genome. To do so, we prepared Pol II ChIP-sequencing datasets from uninfected and infected cells at 4 hours and 12 hours post infection. In doing so, we found that infection induced changes in PAF1 recruitment were also accompanied by changes with Pol II recruitment to the same genes (Figures 4.3A and 4.3B). Metagene profiles of Pol II recruitment across genes that displayed altered PAF1 binding (PAF1-target) against all PAF1-bound genes showed that Pol II was also specifically recruited to PAF1 dependent, but not PAF1 independent genes (Figure 4.3C).

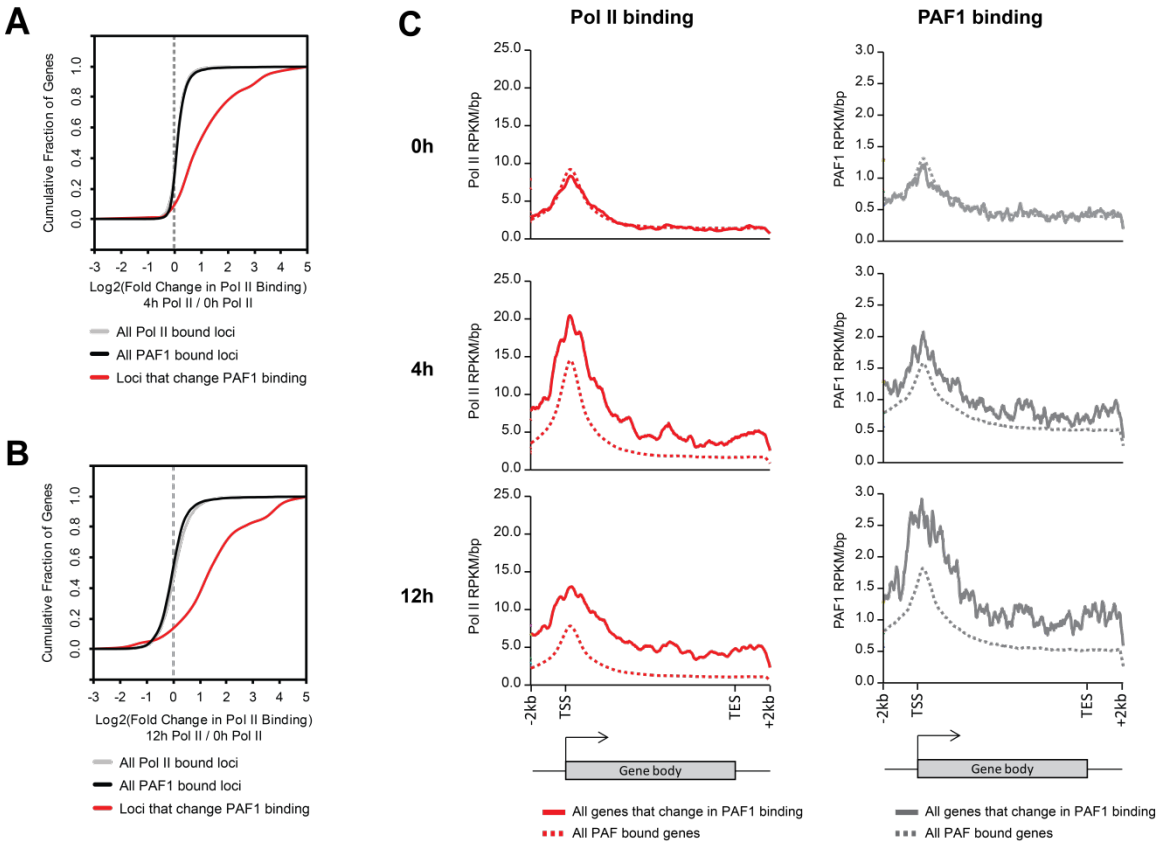


Figure 4.3: Pol II binding correlated to PAF1 binding. A. Cumulative distribution function (CDF) plots of classes of Pol II or PAF1 bound genes at 4 hours post infection. The logged fold change in Pol II binding between infected cells at 4 hours post infection and uninfected cells are shown. Grey line: All Pol II bound genes; Black line: All PAF1 bound genes; Red line: All PAF1 bound genes that display differential recruitment (> 2 fold) during the course of infection. **B.** Same as for A, except displaying logged fold change in Pol II binding between infected cells at 12 hours post infection and uninfected cells. **C.** Metagene profiles of Pol II and PAF1 binding profiles at all genes that show differential recruitment of PAF (Solid lines) or all genes that bind PAF1 during infection (dotted lines). Shown are Pol II (red) and PAF1 (grey) binding profiles at their respective classes in uninfected cells (top panels) and in infected cells at 4 hours (middle panels) and 12 hours (lower panels) post infection.

While changes in abundance of PAF1 and Pol II binding at a gene suggest transcriptional competence, it may not necessarily reflect overall changes in gene expression. Instead, a myriad of other co-transcriptional processes can also impact the

formation of functional transcripts. As such, one approach of determining the impact of PAF1 binding on gene expression changes would be to quantify levels of mature, poly-adenylated transcripts in cells. We thus decided to perform microarray analyses on cells that had been infected with the PR8/ΔNS1virus. When we compared changes in fold expression for PAF1 dependent and PAF1 independent genes, we found that genes that actively recruited PAF1 during infection tended to show an increase in expression levels (Figure 4.4). On the other hand, when we considered all PAF1 bound genes across the genome, the majority of these had more stable expression levels (Figure 4.4) That is, their cumulative distributions were centered around a Log_2(fold change in expression) value of zero.

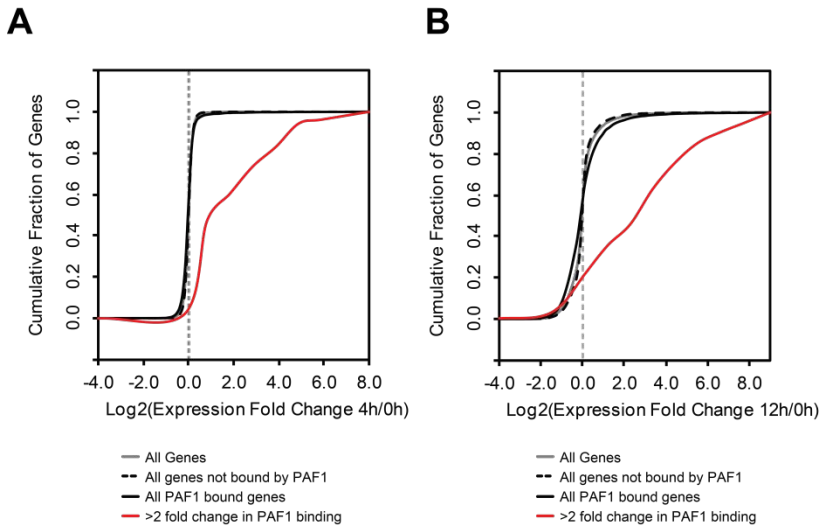


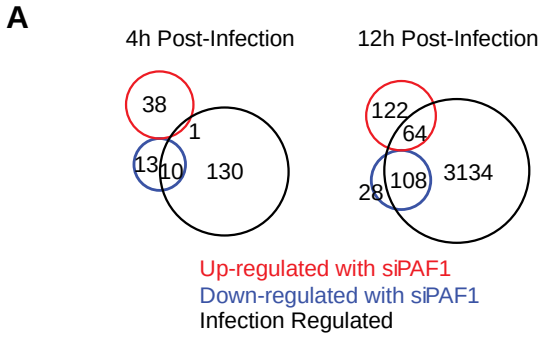
Figure 4.4: Inducible-recruitment of PAF1 genes correlates to an increase in gene expression. A. CDF plots of expression changes within different subsets of PAF1 bound genes at 4 hours post infection. Genes that inducibly recruit PAF1: solid red line, all genes that bind PAF: solid black line, genes that do not bind PAF1: broken black line, and all genes: solid grey line **B.** CDF plots of expression changes within different subsets of PAF1 bound genes at 12 hours post infection.

4.5 PAF1-target genes in infection are sensitive to PAF1 depletion

Our data suggest that the expression of the anti-viral genes is related to the active recruitment of PAF1. This was interesting, in the light of some of our previous observations that PAF1 depletion during infection specifically inhibited the expression of anti-viral genes, but not of housekeeping genes. Active recruitment of PAF1 to these genes might represent a rate limiting step in the activation of the genes that is exacerbated during PAF1 deficiency. This could account for why PAF1, despite its role as a general transcription elongation factor, only impacts the infection regulated genes during infection.

We thus re-did microarray analyses on uninfected and PR8/ΔNS1 infected cells treated either with siRNA against PAF1 (siPAF) or with a control, non-targeting siRNA (siCtrl). Consistent with our previous results, PAF1 deficiency had a limited impact on gene expression in uninfected cells, with 46 genes showing greater than 2 fold change. The expression of housekeeping genes was also stable. Also, as we had shown previously, the impact of PAF1 deficiency was much greater in infected cells, with a total of 367 genes displaying siPAF1 sensitivity over the course of infection (Figure 4.5A). There were two subgroups of siPAF1 sensitive genes. These were (1)genes that were up-regulated upon PAF1 depletion, and (2) those that were down-regulated upon PAF1 depletion. Overall, we found that PAF1 had strong impact on about 7.5% infection regulated genes at 4 hours post infection, and 5.0% of infection regulated genes at 12 hours post infection, indicating that PAF1 depletion only affected a subset of genes (Figure 4.5A). Consistent with what we observed previously, the genes that were

sensitive to PAF1 depletion were enriched with inflammatory genes and anti-viral effector genes (Figure 4.5B).



B

Category	p-value
Cell Death and Survival	1.3E-11
Infectious Disease	1.5E-10
Antimicrobial Response	1.8E-10
Inflammatory Response	1.8E-10
Cellular Development	3.5E-10

Figure 4.5: siPAF1 sensitive genes in infection. A. Overlap of PAF1 sensitive genes and infection regulated genes at 4h and 12h post infection. **B.** Top five functional categories associated with siPAF1 sensitive genes as identified by IPA analysis

We next determined the pattern of PAF1 binding on siPAF1 sensitive genes during infection. To do so, we cross-referenced our microarray data to our PAF1 ChIP-seq data. We were able to map 45-50% of our siPAF1 sensitive genes at 4 and 12 hours post infection to PAF1 bound regions during infection. When we examined PAF1 levels on these genes however, we saw that there was a direct correlation between genes affected by PAF1 depletion and its active recruitment during infection (Figures 4.6A and B middle and right panels). In fact, by 12 hours post infection, PAF1 binding levels at these genes were significantly higher than other PAF1 bound genes genome-wide (Figure 4.6B, right panels). In addition, comparison of these loci against genome wide PAF1 bound loci revealed that the majority of these siPAF1-sensitive loci were initially associated with lower PAF1 binding levels as compared to that of other genes (Figures 4.6A and B, left

panels). In contrast, genes that were up-regulated upon PAF1 depletion showed little or no change in PAF1 binding throughout the infection. The changes in PAF1 binding associated with these genes could not be attributed to a general loss of PAF1 binding at other genes, as the levels of PAF1 on the majority of PAF1 bound loci remained constant throughout infection (Figure 4.6C).

We were concerned that the quantitative aspect of our bioinformatics analysis could have been skewed due to inherent limitations of our peak-calling program. As such, we also examined metagene profiles PAF1 binding at siPAF1 sensitive and all genes in the genome. Consistent with our previous results, we found that siPAF1 sensitive genes were associated with increasing levels of PAF1 during the course of a normal infection (Figure 4.6D). In contrast, metagene profiles genes throughout the whole genome revealed limited change in PAF1 binding during infection. Altogether, these data suggest that siPAF1 sensitive genes tend to actively recruit PAF1 during infection.

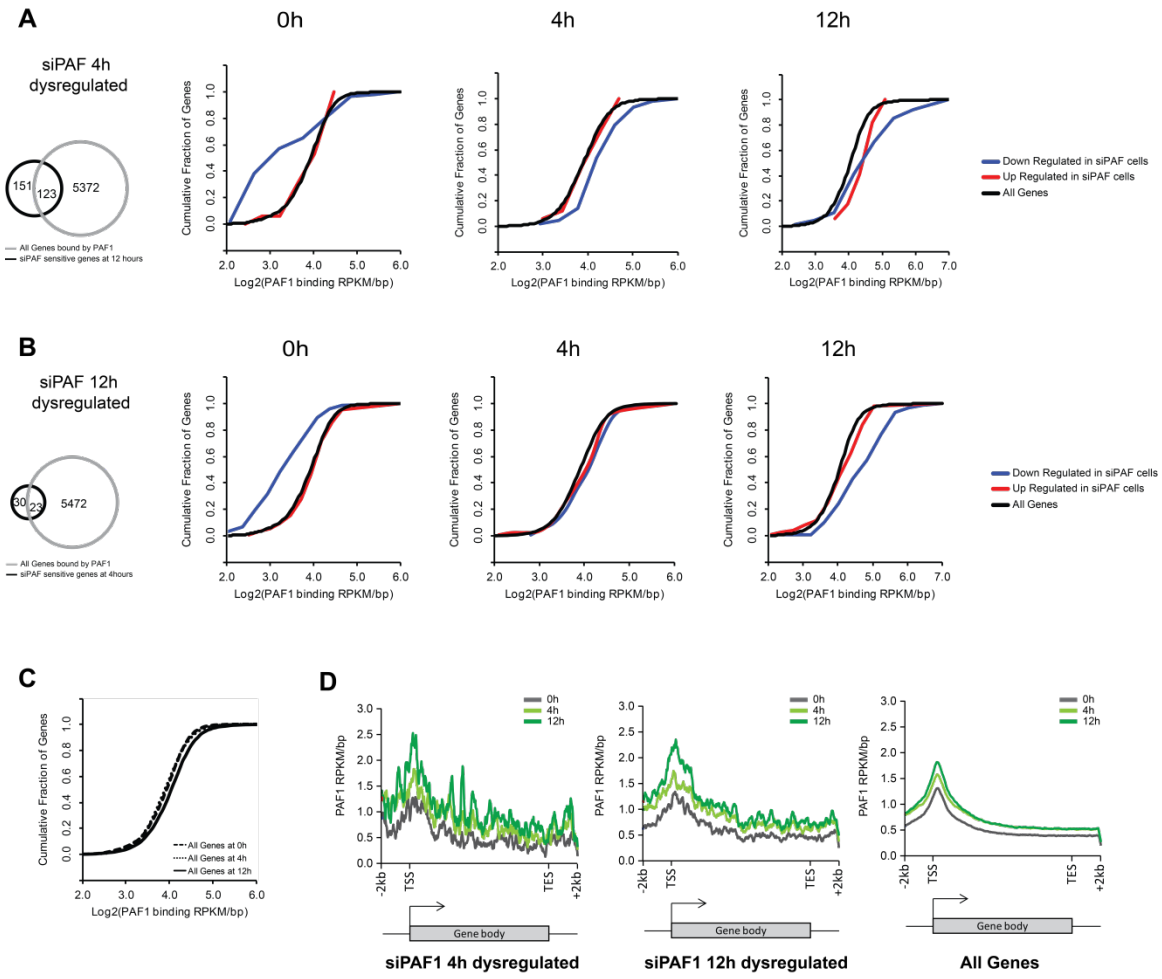


Figure 4.6: PAF1 levels are dynamically regulated in siPAF1 down-regulated genes. **A.** CDF plots PAF1 binding intensity at genes that were identified to be dysregulated by siPAF at 4 hours post infection. Left panel: PAF1 levels at siPAF downregulated (blue), siPAF upregulated (red) and PAF levels at all genes in uninfected (0h) cells; Middle panel: PAF1 levels at the same genes at 4 hours post infection; Right panel: PAF1 levels at the same genes at 12 hours post infection. Note how the amount of PAF associated with the siPAF downregulated genes shift over time. The overlap between siPAF sensitive genes and all PAF1 genes is indicated in the venn diagram. **B.** Same as in A, except comparing PAF1 binding in genes that were dysregulated by siPAF treatment at 12 hours post infection. **C.** Overall levels of PAF1 across the genome remains the same throughout the course of infection. **D.** Metagene profiles showing PAF1 binding in genes that are dysregulated by siPAF at 4 hours (left panel) and 12 hours (middle panel) during infection. The PAF1 binding profiles of all genes is also shown (right panel). For each class of gene, PAF1 binding levels in uninfected cells are shown as solid grey lines, whereas PAF1 binding levels for genes at 4 hours post infection and 12 hours post infection are shown in yellow and green respectively.

4.6 Conclusions

In summary, we have shown that in contrast to the rest of the genome, PAF1 is actively recruited to anti-viral and inflammatory genes. PAF1 recruitment to this subset of genes coincides with the recruitment of Pol II and their expression. We also found that genes that are sensitive to PAF1 depletion during infection have low initial levels of PAF1 bound, but actively recruit PAF1 as the infection progresses. We therefore propose that PAF1 sensitivity is likely to result from an inability to recruit optimal levels of PAF1 to target genes during infection.

CHAPTER 5: DISCUSSION

Starting from an unbiased bioinformatics screen, we were able to show a novel means by which influenza virus can use a histone H3K4-like sequence in its NS1 protein to specifically target inflammatory gene expression. This is achieved through the inhibition of transcription elongation activity of the PAF1C in host cells. In addition, we show that the PAF1C is an important regulator of the inflammatory response, and its active recruitment to anti-viral genes is required for their expression.

5.1 PAF1C and its associated activities are important targets for pathogens

By using chromatinized *in vitro* transcription assays, we were able to show that the binding of NS1 to PAF1 was able to block the transcription elongation activity of PAF1. This corroborated and further explained our *in vivo* genome-wide analysis of transcriptional elongation dynamics (GRO-seq) in infected cells, where we saw a defect in transcription elongation in wild-type NS1 infected cells as compared to deltaPAF1 infected cells. Our results suggest that the NS1 protein specifically targets the transcription elongation activity of the PAF1C in cells.

However, aside from its role as a transcription elongator, the PAF1C has been shown to participate in several other co-transcriptional processes. These include promoting transcription-coupled histone modifications (Hahn et al., 2012; Kim and Roeder, 2009; Wood et al., 2003) as well as regulation of mRNA 3' end processing (Mueller et al., 2004; Nagaike et al., 2011; Nordick et al., 2008; Penheiter et al., 2005; Rozenblatt-Rosen et al., 2009). For instance, the PAF1C has been shown to impact monoubiquitylation of histone H2BK120 in human cells, by promoting recruitment of the

ubiquitin-conjugating enzyme Rad6 and the ubiquitin protein ligase hBre1 to chromatin (Kim and Roeder, 2009; Wood et al., 2003). Rad6/Bre1 mediated monoubiquitylation of H2BK120 facilitates H3K4 tri-methylation and H3K79 di-methylation by the Set1/MLL-1 and Dot1 complexes respectively (Dover et al., 2002; Kim et al., 2009; Nakanishi et al., 2009; Sun and Allis, 2002). In turn, these modifications can influence the recruitment and/or activities that further impact chromatin structure and transcriptional competence of the gene.

Given the multiple roles that the PAF1C plays in regulating transcription and in enhancing inflammatory gene transcription, it is no wonder that the complex and its associated activities are subject to manipulation by several other pathogens besides the Influenza A virus. Indeed, the human immunodeficiency virus (HIV) TAT protein was previously shown to recruit the PAF1 complex, along with P-TEFb and other elongation factors (specifically the super elongation complex, comprising of AFF4, AF9, ELL and ENL) (Sobhian et al., 2010) to the virus promoter during infection. This was thought to promote the formation of a permissive chromatin environment and allow viral replication. In further support of this, independently conducted studies have shown that PAF1 over-expression results in overall reduced levels of pro-viral integration during HIV infection, highlighting the importance of PAF1 expression in the HIV viral life cycle (Liu et al., 2011).

Interestingly, the human adenovirus (HAdV) has also been shown to require PAF1C activity to replicate. This is mediated via the HAdV virulence factor E1A, which binds to and inhibits Bre1 protein, which, as mentioned previously, also interacts with the PAF1C. During infection, E1A recruits the PAF1C to the viral genome via Bre1, thus

promoting viral early gene expression (Fonseca et al., 2013). In support of this, knockdown of PAF1 results in reduced viral early gene transcription. The interaction of E1A to Bre1 was also shown to inhibit the expression of IFN and ISGs, suggesting that Rad6/Bre1 mediated ubiquitylation of H2BK120 was important for the expression of these genes (Fonseca et al., 2012). Given that Rad6/Bre1 activity is also facilitated by the PAF1 complex, these data support the notion that the PAF1C and its related activities are key regulators of the inflammatory response.

5.2 Differential Susceptibility of stress-response genes to PAF1 depletion

We were interested to note that despite the general association of PAF1 with Pol II genome-wide, only a subset of genes within the cell are sensitive to PAF1 depletion. This phenomenon was not restricted to our experimental system. In yeast, PAF1 knockout (PAF1Δ) cells behave normally under homeostatic conditions, but are hypersensitive to a variety of stress-inducing conditions (Betz et al., 2002; Costa and Arndt, 2000; Kim and Levin, 2011; Squazzo et al., 2002). Altogether, these data suggest that the PAF1C serves a function in stress-induced gene expression.

Our studies imply that genes that are most sensitive to PAF1 depletion are genes that have low initial levels of PAF1 bound and require active recruitment of PAF1 to become active. Reduced or inhibited PAF1 recruitment to these genes by siRNA-mediated depletion of PAF1 thus inhibits gene activation. Many of the stress response genes can impact normal biological processes of the cell, such as protein folding, translation or pre-mRNA splicing. As such, these genes are tightly regulated and silenced under homeostatic conditions, since their constitutive activation is could be detrimental to

the cell. Given the central role of PAF1 in transcription, its recruitment is likely to affect several regulatory steps in the activation of these genes.

PAF1 has been implicated in several events that control the efficiency of transcription elongation of its target genes. As mentioned earlier, an important activity of the PAF1 complex is the promotion of histone H2B ubiquitylation, which has been shown to be critical for the transcription of several inducible genes (including GAL1, SUC2 and PHO5) in yeast (Kao et al., 2004). In the absence of ubiquitylated H2b, transcription of these genes is severely impaired. Also, as mentioned previously, HAdV mediated inhibition of the hBre1 complex results in reduced H2B ubiquitylation and an impaired IFN response in infected cells (Fonseca et al., 2012). H2B ubiquitylation may also affect promoter clearance and release of Pol II into productive elongation. In yeast, H2B de-ubiquitylation by Ubp8 is required to recruit Ctk kinase complex to the gene(Henry et al., 2003). In turn, Ctk phosphorylation of the serine 2 in the Pol II CTD facilitates Pol II entry into productive elongation (Zhou et al., 2009). Given that the PAF1C facilitates the recruitment and activation of the Rad6/Bre1 complex to ubiquitinate histone H2B, it may be that this activity contributes to the reliance of stress induced genes on PAF1C recruitment. It will thus be interesting to investigate how Rad6/Bre1 and H2B ubiquitylation is affected during infection.

Besides its interactions with histone modifying enzymes such as Rad6/Bre1, the PAF1C has also been shown to interact with a number of transcription elongation complexes, including DSIF, FACT, TFIIS and SEC (Super Elongator complex) (Chen et al., 2009; Dawson et al., 2011; Kim et al., 2010; Qiu et al., 2006; Squazzo et al., 2002). PAF1C has been shown by several groups to stimulate transcription of target genes

through cooperative interactions with these complexes. These complexes have all been implicated in different regulatory steps during inducible transcription. In fact, loss of PAF1 results in reduced association of DSIF and FACT on chromatin (Mueller et al., 2004; Pruneski et al., 2011), suggesting that it is required for these factors to maintain association with their targets and promote transcription.

Finally, the PAF1 complex has been shown to directly stimulate transcription elongation rates on chromatinized template (Kim et al., 2010). Indeed, our results suggest that it is this PAF1 dependent activity that is targeted by NS1 during the suppression of the inflammatory response. This implies that transcription elongation efficiency of stress response genes is in itself important for their full induction. This has been similarly suggested in other inducible transcription systems (Danko et al., 2013; Hah et al., 2011). Further support for the role of transcriptional elongation in the regulation of the stress response has been shown in studies of a PAF1C independent transcription elongation factor Elongin A. Elongin A is a component of a multisubunit transcription elongation complex comprising of Elongin A, Elongin B and Elongin C. Like the PAF1 protein, previous studies have shown that Elongin A is not generally required for transcription *in vivo* under homeostatic conditions (Gerber et al., 2005; Kawauchi et al., 2013). However, during the induction of the heat shock response, induction of the heat shock responsive genes ATF3 and HSP70 were abrogated. Taken together, these studies strongly support the notion that transcription elongation is an important component for the efficient induction of stress response genes. An important implication of these studies is that inhibition of transcription elongation activity could allow for a targeted way of manipulating stress response genes. Given that many diseases are manifested by over-

activation of stress responses, these findings could have important implications in the development of therapeutics aimed at balancing the out of control stress response.

5.3 Mechanisms of PAF1C recruitment to target genes

Our data show that the active recruitment of PAF1 to a subset of anti-viral genes coincides with and is critical for their expression during infection. This is likely to contribute to their sensitivity to PAF1 depletion during infection. However, the mechanisms by which PAF1C is recruited to its target genes on chromatin is not well understood in mammalian cells. In yeast, recruitment of the yeast PAF1C (yPAF1C) to chromatin is mediated primarily by its yRTF1 and yCDC73 subunits (Amrich et al., 2012; Warner et al., 2007). yRTF1 interaction with chromatin is thought to be mediated by a conserved Plus3 domain. Loss of this domain in yRTF1 results in the impaired chromatin association of yPAF1C (Warner et al., 2007). In addition, yRTF1 has been shown to play an important role in mediating histone H2B ubiquitylation, and a small histone modification domain present on the protein is sufficient to substitute the entire yPAF1C for promoting H2B ubiquitylation in the absence of PAF1(Piro et al., 2012). On the other hand, yCDC73 was also shown to be important yPAF1C recruitment to genes. This is thought to be mediated by a Ras-like domain in the C-terminal of yCDC73. Truncation of almost the entire domain in yCDC73 impairs yPAF1C recruitment to chromatin without disrupting the integrity of yPAF1C(Amrich et al., 2012). Independent studies confirmed a role of yCDC73 in binding to the phophorylated C-terminal domain (CTD) of elongating Pol II and also to the phosphorylated C-terminal region (CTR) of

yeast Spt5 (subunit of the yeast Spt4/Spt5 complex, and yeast ortholog of the mammalian DSIF) (Qiu et al., 2012; Qiu et al., 2006).

Another group of candidates that have been implicated in PAF1C recruitment to chromatin is the bromodomain and ET domain (BET) family of proteins. The BET proteins are a family of reader proteins that bind to acetylated histones. Members of the family, which include BRD2, BRD3, BRD4 and BRDT, share similar structural properties, each comprising of two tandem bromodomain modules, an extra-terminal domain and a C-terminal domain (Belkina and Denis, 2012). Interestingly, recent proteomic screens have identified that the PAF1C may interact and function cooperatively with members of this family (Dawson et al., 2011). In fact, these studies showed that inhibition of BET binding activity by I-BET, resulted in reduced PAF1 recruitment to chromatin (Dawson et al., 2011). In addition to this, similar to PAF1 depletion, inhibition of BET protein recruitment to chromatin results in the selective repression of inflammatory genes during lipopolysaccharide (LPS) stimulation of bone-marrow derived macrophages (Hargreaves et al., 2009; Nicodeme et al., 2010). These studies, taken together with our findings that PAF1 recruitment is essential for the proper induction of the anti-viral response, suggest that the BET proteins could be critical for maintaining optimal levels of PAF1C on chromatin. It remains to be seen how the BET proteins and PAF1C interact to coordinate the transcriptional response in infection.

Finally, a recent study revealed that histone H3 asymmetrically di-methylated on arginine 17 (H3R17) could promote the recruitment of PAF1C to chromatin(Wu and Xu, 2012). Consistent with our results, this study demonstrated that an unmodified histone H3 tail is sufficient to bind PAF1C, but that this interaction is enhanced upon methylation of

H3R17. Loss of the CARM1 methyltransferase, the primary mediator of H3R17 methylation, results in greatly reduced PAF1 recruitment to chromatin. In contrast, loss of depletion of PAF1C components had no impact on H3R17me2a abundance, indicating that CARM1 and H3R17me2a act upstream of PAF1C. Taken together with our findings on PAF1, this result not only suggests a potential cross talk between H3K4 and H3R17, but also strongly implies a role for CARM1 in the control of the inflammatory response. Indeed, CARM1 has previously been implicated as a co-activator for NF-kB dependent genes, and was shown to enhance their activities during stimulation (Miao et al., 2006). Further investigation will be required to uncover the role of the CARM1 and the PAF1C in regulating the inflammatory response.

5.4 Histone mimicry is a viable strategy for pathogens to manipulate host processes

This study has demonstrated that histone mimicry can be a viable strategy for pathogens to modulate host function. We were also able to confirm that the histone H3 protein can interact with PAF1C. Altogether, these data support the idea that the histone H3 motifs can function as discrete informational units in signaling. It will be interesting to investigate if other viral pathogens that we found in our *in silico* screen also carry functional histone-motifs.

A key feature of the histone tails is their ability to switch between different functional states via post-translational modifications (see Introduction). Indeed, we show that PAF1 binding to the histone H3 tails can be abrogated by acetylation, but not methylation. The same appears to be true for the NS1 histone-like sequence, although we

were regrettably not able to fully investigate the functional consequences of NS1 or Histone H3 acetylation and PAF1 binding. However, it may be that in addition to inhibiting PAF1 activity to suppress the immune response, NS1 could also use the PAF1C to locate its other binding partners. Indeed, the PAF1C has been shown to bind to another known NS1 target, the cleavage and polyadenylation factors complex (CPSF) (Rozenblatt-Rosen et al., 2009) through its CDC73 subunit.

Modification of NS1 could thus play a role in "re-purposing" the NS1 protein, and allow it to switch between binding partners during infection. For example, both CHD1 and PAF1C, which have been shown to interact with each other (Simic et al., 2003), were able to bind to the NS1 histone-like sequence, albeit in a modification-dependent ways. Moreover, there is increased interaction of the NS1 protein with host CHD1 upon methylation of the NS1 histone-like sequence. Since CHD1 has also previously been shown to interact with spliceosomal components and impact efficacy of pre-mRNA splicing (Sims et al., 2007) (in addition to its role as a chromatin remodeler (Tran et al., 2000)), this could potentially be a way for the virus to gain access to the spliceosome and impact pre-mRNA splicing in the host cell (Lu et al., 1994; Qiu et al., 1995; Wang and Krug, 1998). NS1 interaction with the host splicing machinery may also play an important role in regulating viral mRNA splicing, which is also important for the viral lifecycle (Chua et al., 2013).

5.5 Short linear motifs occurring in the NS1 C-terminal domain are virulence factors

The NS1 C-terminal tail domain has previously been implicated as a virulence factor for some avian influenza A viral isolates. In fact, several studies have shown that the majority of NS1 sequences derived from these viruses carry PDZ ligand (PL) motifs (consensus: X-[S/T]-X-V_{COOH}; where X represents any amino acid, and COOH represents the C-terminal) in their C-terminal domain (Figure 5.1) (Obenauer et al., 2006). PDZ domains are protein-protein recognition motifs that are found widely in proteins involved in cell-signaling. These proteins specifically recognize and bind to short 4 to 5 amino acid long PL motifs, which are almost always found at the C-terminus of the protein. Avian PL motifs (including EPEV, ESEV) are not commonly observed in human isolates of virus, but have been strictly associated with several highly pathogenic human isolates. In particular, NS1 sequences from the H5N1 1997 and H5N1 2003 human pandemic viruses were shown to contain avian-like C-terminal PL domains ESEV and EPEV respectively (Jackson et al., 2008; Obenauer et al., 2006; Thomas et al., 2011). These PL motifs were necessary for the interaction of the virus with PDZ proteins such as Dlg, scribble (Golebiewski et al., 2011), Magi-I (Kumar et al., 2012). In particular, H5N1 NS1 PL-mediated binding to scribble and Dlg1 was required to protect infected cells from apoptosis, thus contributing to viral virulence (Golebiewski et al., 2011).

On the other hand, common-circulating, low pathogenic human viruses predominantly contain PL motif sequences RSKV (H3N2 isolates), and RSEV (H1N1 isolates). Despite being canonical PLs, these motifs do not appear to bind well to known PDZ domain containing proteins (Obenauer et al., 2006; Thomas et al., 2011). In fact,

viral virulence can be increased if these "human" PL motifs are replaced with those derived from avian isolates. Despite this, overall loss of the NS1 C-terminal tail from low-pathogenic viruses still results in viral attenuation(Jackson et al., 2008). This suggests that the C-terminal domain of NS1 is an important contributor to the overall virulence of the virus, although its specific function remains unknown. In this study, we show that the H3N2 virus utilizes its NS1 C-terminal domain to interact with host PAF1C instead. This is mediated by a histone H3-like motif A-R-S/T-K also present in the C-terminal domain. NS1 interaction with PAF1 was necessary for the virus to repress the host immune response, and replicate well. Altogether, these studies confirm the role of the role of the Influenza A NS1 C-terminal domain as an important virulence factor that as likely evolved to interact in a species-specific way with pivotal regulators of the cell response.

5.6 Influenza Adaptation and Virulence

It is interesting to note that the histone H3-like sequence is predominantly found in the low-pathogenic human influenza viruses. Historically, low virulence has often been argued as a way for parasites to avoid overexploitation of the host resources, and improve transmissibility (Frank, 1996). This theory postulates that given enough time, pathogens would eventually become avirulent and achieve symbiosis with their hosts. In the light of this, it is tempting to speculate that the interaction of the H3N2 NS1 constitutes an adaptation of the virus to maintain a virus pool circulating in the human population.

However, the truth is that virulence of the influenza virus is a multi-genic trait that is not dependent on NS1 alone (Fukuyama and Kawaoka, 2011; Tscherne and Garcia-

Sastre, 2011; Wasilenko et al., 2008). Several studies have indicated that influenza viral proteins are functionally linked, and exert epistatic effects on one another (Kaverin et al., 2000; Twu et al., 2007). In fact, NS1 protein function has been shown to be tightly linked to the identity of viral polymerase genes(Kuo and Krug, 2009). Notably, artificially engineering reassortant viruses, where the cognate NS segment is replaced with a NS gene from more virulent strains of virus, has been shown to have minimal effects on viral pathogenicity (Kuo and Krug, 2009; Sarmento et al., 2010; Shelton et al., 2012). Furthermore, large scale sequencing analysis of avian influenza isolates has also indicated that many of the core influenza protein co-segregate with one another (Obenauer et al., 2006).

The lower pathogenicity of the human influenza viruses (including H3N2 and H1N1) also do not necessarily indicate a trend towards avirulence and symbiosis of the Influenza A virus within human populations. This model assumes that viral transmission is a function of host mortality (i.e. viruses that kill their hosts quickly have a poorer chance of survival), while ignoring other factors, such as the geographical distribution of host populations, or potential competition between related strains of the same virus. In fact, several groups have argued that local distributions and intrinsic susceptibility of host populations are more important factors for determining viral virulence (Boots and Sasaki, 1999; Frank and Schmid-Hempel, 2008; Palese and Wang, 2011; Pfennig, 2001; Wild et al., 2009). The correlation between NS1-PAF1 interaction and poorer viral virulence could indeed be an evolutionary step towards achieving optimal viral fitness, but must ultimately be considered in the context of NS1 interactions with other viral and host proteins, as well as pathogen specialization within host populations.

BIBLIOGRAPHY

Abada, R., Dreyfuss-Grossman, T., Herman-Bachinsky, Y., Geva, H., Masa, S.R., and Sarid, R. (2008). SIAH-1 interacts with the Kaposi's sarcoma-associated herpesvirus-encoded ORF45 protein and promotes its ubiquitylation and proteasomal degradation. J Virol *82*, 2230-2240.

Agalioti, T., Chen, G.Y., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Agalioti, T., Lomvardas, S., Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. Cell *103*, 667-678.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell *124*, 783-801.

Amrich, C.G., Davis, C.P., Rogal, W.P., Shirra, M.K., Heroux, A., Gardner, R.G., Arndt, K.M., and VanDemark, A.P. (2012). Cdc73 subunit of Paf1 complex contains C-terminal Ras-like domain that promotes association of Paf1 complex with chromatin. J Biol Chem *287*, 10863-10875.

Ansieau, S., and Leutz, A. (2002). The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif. J Biol Chem *277*, 4906-4910.

Arbibe, L., Kim, D.W., Batsche, E., Pedron, T., Mateescu, B., Muchardt, C., Parsot, C., and Sansonetti, P.J. (2007). An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. Nat Immunol *8*, 47-56.

Baccala, R., Gonzalez-Quintial, R., Lawson, B.R., Stern, M.E., Kono, D.H., Beutler, B., and Theofilopoulos, A.N. (2009). Sensors of the innate immune system: their mode of action. Nat Rev Rheumatol *5*, 448-456.

Badeaux, A.I., and Shi, Y. (2013). Emerging roles for chromatin as a signal integration and storage platform. Nat Rev Mol Cell Biol *14*, 211-224.

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.

Becker, P.B., and Horz, W. (2002). ATP-dependent nucleosome remodeling. Annu Rev Biochem *71*, 247-273.

Belkina, A.C., and Denis, G.V. (2012). BET domain co-regulators in obesity, inflammation and cancer. Nat Rev Cancer *12*, 465-477.

Betz, J.L., Chang, M., Washburn, T.M., Porter, S.E., Mueller, C.L., and Jaehning, J.A. (2002). Phenotypic analysis of Paf1/RNA polymerase II complex mutations reveals

connections to cell cycle regulation, protein synthesis, and lipid and nucleic acid metabolism. Mol Genet Genomics *268*, 272-285.

Bhaumik, S.R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol *14*, 1008-1016.

Boots, M., and Sasaki, A. (1999). 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. Proc Biol Sci *266*, 1933-1938.

Bornholdt, Z.A., and Prasad, B.V. (2006). X-ray structure of influenza virus NS1 effector domain. Nat Struct Mol Biol *13*, 559-560.

Bouvier, N.M., and Palese, P. (2008). The biology of influenza viruses. Vaccine *26 Suppl 4*, D49-53.

Boyd, J.M., Subramanian, T., Schaeper, U., La Regina, M., Bayley, S., and Chinnadurai, G. (1993). A region in the C-terminus of adenovirus 2/5 E1a protein is required for association with a cellular phosphoprotein and important for the negative modulation of T24-ras mediated transformation, tumorigenesis and metastasis. EMBO J *12*, 469-478.

Brower-Toland, B., Wacker, D.A., Fulbright, R.M., Lis, J.T., Kraus, W.L., and Wang, M.D. (2005). Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. J Mol Biol *346*, 135-146.

Bruton, R.K., Pelka, P., Mapp, K.L., Fonseca, G.J., Torchia, J., Turnell, A.S., Mymryk, J.S., and Grand, R.J. (2008). Identification of a second CtBP binding site in adenovirus type 5 E1A conserved region 3. J Virol *82*, 8476-8486.

Carvalho, T., Seeler, J.S., Ohman, K., Jordan, P., Pettersson, U., Akusjarvi, G., Carmo-Fonseca, M., and Dejean, A. (1995). Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. J Cell Biol *131*, 45-56.

Chang, Y., Sun, L., Kokura, K., Horton, J.R., Fukuda, M., Espejo, A., Izumi, V., Koomen, J.M., Bedford, M.T., Zhang, X.*, et al.* (2011). MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. Nat Commun *2*, 533.

Chen, Y., Yamaguchi, Y., Tsugeno, Y., Yamamoto, J., Yamada, T., Nakamura, M., Hisatake, K., and Handa, H. (2009). DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. Genes Dev *23*, 2765-2777.

Chien, C.Y., Tejero, R., Huang, Y., Zimmerman, D.E., Rios, C.B., Krug, R.M., and Montelione, G.T. (1997). A novel RNA-binding motif in influenza A virus non-structural protein 1. Nat Struct Biol *4*, 891-895.

Chin, H.G., Esteve, P.O., Pradhan, M., Benner, J., Patnaik, D., Carey, M.F., and Pradhan, S. (2007). Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. Nucleic Acids Res *35*, 7313-7323.

Chua, M.A., Schmid, S., Perez, J.T., Langlois, R.A., and Tenoever, B.R. (2013). Influenza A virus utilizes suboptimal splicing to coordinate the timing of infection. Cell Rep *3*, 23-29.

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.

Corona, D.F., Clapier, C.R., Becker, P.B., and Tamkun, J.W. (2002). Modulation of ISWI function by site-specific histone acetylation. EMBO Rep *3*, 242-247.

Costa, P.J., and Arndt, K.M. (2000). Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation. Genetics *156*, 535-547.

Craig, H.M., Pandori, M.W., and Guatelli, J.C. (1998). Interaction of HIV-1 Nef with the cellular dileucine-based sorting pathway is required for CD4 down-regulation and optimal viral infectivity. Proc Natl Acad Sci U S A *95*, 11229-11234.

Danko, C.G., Hah, N., Luo, X., Martins, A.L., Core, L., Lis, J.T., Siepel, A., and Kraus, W.L. (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol Cell *50*, 212-222.

Daugherty, M.D., and Malik, H.S. (2012). Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet *46*, 677-700.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol *319*, 1097-1113.

Davey, N.E., Trave, G., and Gibson, T.J. (2011). How viruses hijack cell regulation. Trends Biochem Sci *36*, 159-169.

Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H., and Gibson, T.J. (2012). Attributes of short linear motifs. Mol Biosyst *8*, 268-281.

Dawson, M.A., Prinjha, R.K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.I., Robson, S.C., Chung, C.W., Hopf, C., Savitski, M.M., *et al.* (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature *478*, 529-533.

Deng, L.W., de la Fuente, C., Fu, P., Wang, L., Donnelly, R., Wade, J.D., Lambert, P., Li, H., Lee, C.G., and Kashanchi, F. (2000). Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones. Virology *277*, 278-295.

Deng, L.W., Wang, D., de la Fuente, C., Wang, L., Li, H., Lee, C.G., Donnelly, R., Wade, J.D., Lambert, P., and Kashanchi, F. (2001). Enhancement of the p300 HAT activity by HIV-1 Tat on chromatin DNA. Virology *289*, 312-326.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., and Gibson, T.J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci *13*, 6580-6603.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. Journal of Biological Chemistry *277*, 28368-28371.

Dutnall, R.N., and Ramakrishnan, V. (1997). Twists and turns of the nucleosome: tails without ends. Structure *5*, 1255-1259.

Eberharter, A., and Becker, P.B. (2004). ATP-dependent nucleosome remodelling: factors and functions. J Cell Sci *117*, 3707-3711.

Egorov, A., Brandt, S., Sereinig, S., Romanova, J., Ferko, B., Katinger, D., Grassauer, A., Alexandrova, G., Katinger, H., and Muster, T. (1998). Transfectant influenza A viruses with long deletions in the NS1 protein grow efficiently in Vero cells. J Virol *72*, 6437-6441.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Elde, N.C., and Malik, H.S. (2009). The evolutionary conundrum of pathogen mimicry. Nat Rev Microbiol *7*, 787-797.

Fang, T.C., Schaefer, U., Mecklenbrauker, I., Stienen, A., Dewell, S., Chen, M.S., Rioja, I., Parravicini, V., Prinjha, R.K., Chandwani, R., *et al.* (2012). Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. J Exp Med *209*, 661-669.

Fernandez-Capetillo, O., Chen, H.T., Celeste, A., Ward, I., Romanienko, P.J., Morales, J.C., Naka, K., Xia, Z., Camerini-Otero, R.D., Motoyama, N., *et al.* (2002). DNA damage-induced G2-M checkpoint activation by histone H2AX and 53BP1. Nat Cell Biol *4*, 993-997.

Ferreon, J.C., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2009). Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. Proc Natl Acad Sci U S A *106*, 13260-13265.

Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., *et al.* (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. Cell *149*, 214-231.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O., Morse, E.M., Keates, T., Hickman, T.T., Felletar, I., *et al.* (2010). Selective inhibition of BET bromodomains. Nature *468*, 1067-1073.

Fischle, W., Wang, Y., and Allis, C.D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Fodor, E., Devenish, L., Engelhardt, O.G., Palese, P., Brownlee, G.G., and Garcia-Sastre, A. (1999). Rescue of influenza A virus from recombinant DNA. J Virol *73*, 9679-9682.

Fonseca, G.J., Cohen, M.J., Nichols, A.C., Barrett, J.W., and Mymryk, J.S. (2013). Viral Retasking of hBre1/RNF20 to Recruit hPaf1 for Transcriptional Activation. PLoS Pathog *9*, e1003411.

Fonseca, G.J., Thillainadesan, G., Yousef, A.F., Ablack, J.N., Mossman, K.L., Torchia, J., and Mymryk, J.S. (2012). Adenovirus Evasion of Interferon-Mediated Innate Immunity by Direct Antagonism of a Cellular Histone Posttranslational Modification. Cell Host Microbe *11*, 597-606.

Frank, S.A. (1996). Models of parasite virulence. Q Rev Biol *71*, 37-78.

Frank, S.A., and Schmid-Hempel, P. (2008). Mechanisms of pathogenesis and the evolution of parasite virulence. J Evol Biol *21*, 396-404.

Fukuyama, S., and Kawaoka, Y. (2011). The pathogenesis of influenza virus infections: the contributions of virus and host factors. Curr Opin Immunol *23*, 481-486.

Garcia-Sastre, A., Egorov, A., Matassov, D., Brandt, S., Levy, D.E., Durbin, J.E., Palese, P., and Muster, T. (1998). Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. Virology *252*, 324-330.

Gerber, M., Tenney, K., Conaway, J.W., Conaway, R.C., Eissenberg, J.C., and Shilatifard, A. (2005). Regulation of heat shock gene expression by RNA polymerase II elongation factor, Elongin A. J Biol Chem *280*, 4017-4020.

Gires, O., Kohlhuber, F., Kilger, E., Baumann, M., Kieser, A., Kaiser, C., Zeidler, R., Scheffer, B., Ueffing, M., and Hammerschmidt, W. (1999). Latent membrane protein 1 of Epstein-Barr virus interacts with JAK3 and activates STAT proteins. EMBO J *18*, 3064-3073.

Glenn, J.S., Watson, J.A., Havel, C.M., and White, J.M. (1992). Identification of a prenylation site in delta virus large antigen. Science *256*, 1331-1333.

Golebiewski, L., Liu, H., Javier, R.T., and Rice, A.P. (2011). The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions. J Virol *85*, 10639-10648.

Hacques, M.F., Muller, S., De Murcia, G., Van Regenmortel, M.H., and Marion, C. (1990). Use of an immobilized enzyme and specific antibodies to analyse the accessibility and role of histone tails in chromatin structure. Biochem Biophys Res Commun *168*, 637-643.

Hah, N., Danko, C.G., Core, L., Waterfall, J.J., Siepel, A., Lis, J.T., and Kraus, W.L. (2011). A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell *145*, 622-634.

Hahn, M.A., Dickson, K.A., Jackson, S., Clarkson, A., Gill, A.J., and Marsh, D.J. (2012). The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. Hum Mol Genet *21*, 559-568.

Hale, B.G., Barclay, W.S., Randall, R.E., and Russell, R.J. (2008a). Structure of an avian influenza A virus NS1 protein effector domain. Virology *378*, 1-5.

Hale, B.G., Randall, R.E., Ortin, J., and Jackson, D. (2008b). The multifunctional NS1 protein of influenza A viruses. J Gen Virol *89*, 2359-2376.

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of inducible gene expression by signal-dependent transcriptional elongation. Cell *138*, 129-145.

Harty, R.N., Brown, M.E., Wang, G., Huibregtse, J., and Hayes, F.P. (2000). A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. Proc Natl Acad Sci U S A *97*, 13871-13876.

Hatada, E., and Fukuda, R. (1992). Binding of influenza A virus NS1 protein to dsRNA in vitro. J Gen Virol *73 (Pt 12)*, 3325-3329.

Henry, K.W., Wyce, A., Lo, W.S., Duggan, L.J., Emre, N.C., Kao, C.F., Pillus, L., Shilatifard, A., Osley, M.A., and Berger, S.L. (2003). Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev *17*, 2648-2663.

Herrera, F.J., and Triezenberg, S.J. (2004). VP16-dependent association of chromatin-modifying coactivators and underrepresentation of histones at immediate-early gene promoters during herpes simplex virus infection. Journal of Virology *78*, 9689-9696.

Huang, R.C., and Bonner, J. (1962). Histone, a suppressor of chromosomal RNA synthesis. Proc Natl Acad Sci U S A *48*, 1216-1222.

Imbalzano, A.N., Kwon, H., Green, M.R., and Kingston, R.E. (1994). Facilitated Binding of Tata-Binding Protein to Nucleosomal DNA. Nature *370*, 481-485.

Ito, T., Levenstein, M.E., Fyodorov, D.V., Kutach, A.K., Kobayashi, R., and Kadonaga, J.T. (1999). ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. Genes Dev *13*, 1529-1539.

Izumi, K.M., and Kieff, E.D. (1997). The Epstein-Barr virus oncogene product latent membrane protein 1 engages the tumor necrosis factor receptor-associated death domain protein to mediate B lymphocyte growth transformation and activate NF-kappaB. Proc Natl Acad Sci U S A *94*, 12592-12597.

Jackson, D., Hossain, M.J., Hickman, D., Perez, D.R., and Lamb, R.A. (2008). A new influenza virus virulence determinant: the NS1 protein four C-terminal residues modulate pathogenicity. Proc Natl Acad Sci U S A *105*, 4381-4386.

Jaehning, J.A. (2010). The Paf1 complex: platform or player in RNA polymerase II transcription? Biochim Biophys Acta *1799*, 379-388.

Kao, C.F., Hillyer, C., Tsukuda, T., Henry, K., Berger, S., and Osley, M.A. (2004). Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. Gene Dev *18*, 184-195.

Kaverin, N.V., Matrosovich, M.N., Gambaryan, A.S., Rudneva, I.A., Shilov, A.A., Varich, N.L., Makarova, N.V., Kropotkina, E.A., and Sinitsin, B.V. (2000). Intergenic HA-NA interactions in influenza A virus: postreassortment substitutions of charged amino acid in the hemagglutinin of different subtypes. Virus Res *66*, 123-129.

Kawai, T., and Akira, S. (2007). Antiviral signaling through pattern recognition receptors. J Biochem *141*, 137-145.

Kawauchi, J., Inoue, M., Fukuda, M., Uchida, Y., Yasukawa, T., Conaway, R.C., Conaway, J.W., Aso, T., and Kitajima, S. (2013). Transcriptional properties of mammalian Elongin A and its role in stress response. J Biol Chem.

Kelly, A.E., Ghenoiu, C., Xue, J.Z., Zierhut, C., Kimura, H., and Funabiki, H. (2010). Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. Science *330*, 235-239.

Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat Biotechnol *26*, 1351-1359.

Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T., *et al.* (1999). HIV-1 Tat transcriptional activity is regulated by acetylation. Embo Journal *18*, 6106-6118.

Kim, J., Guermah, M., McGinty, R.K., Lee, J.S., Tang, Z.Y., Milne, T.A., Shilatifard, A., Muir, T.W., and Roeder, R.G. (2009). RAD6-Mediated Transcription-Coupled H2B Ubiquitylation Directly Stimulates H3K4 Methylation in Human Cells. Cell *137*, 459-471.

Kim, J., Guermah, M., and Roeder, R.G. (2010). The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. Cell *140*, 491-503.

Kim, J., and Roeder, R.G. (2009). Direct Bre1-Paf1 Complex Interactions and RING Finger-independent Bre1-Rad6 Interactions Mediate Histone H2B Ubiquitylation in Yeast. Journal of Biological Chemistry *284*, 20582-20592.

Kim, K.Y., and Levin, D.E. (2011). Mpk1 MAPK Association with the Paf1 Complex Blocks Sen1-Mediated Premature Transcription Termination. Cell *144*, 745-756.

Kim, S.J., Ding, W., Albrecht, B., Green, P.L., and Lairmore, M.D. (2003). A conserved calcineurin-binding motif in human T lymphotropic virus type 1 p12I functions to modulate nuclear factor of activated T cell activation. J Biol Chem *278*, 15550-15557.

Kochs, G., Garcia-Sastre, A., and Martinez-Sobrido, L. (2007). Multiple anti-interferon actions of the influenza A virus NS1 protein. J Virol *81*, 7011-7021.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kumar, M., Liu, H., and Rice, A.P. (2012). Regulation of interferon-beta by MAGI-1 and its interaction with influenza A virus NS1 protein with ESEV PBM. PLoS One *7*, e41251.

Kuo, R.L., and Krug, R.M. (2009). Influenza A Virus Polymerase Is an Integral Component of the CPSF30-NS1A Protein Complex in Infected Cells. Journal of Virology *83*, 1611-1616.

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 Biol *10*, R25.

Laybourn, P.J., and Kadonaga, J.T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science *254*, 238-245.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat Protoc *1*, 729-748.

Lewis, P.W., Muller, M.M., Koletsky, M.S., Cordero, F., Lin, S., Banaszynski, L.A., Garcia, B.A., Muir, T.W., Becher, O.J., and Allis, C.D. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science *340*, 857-861.

Li, H., Fischle, W., Wang, W., Duncan, E.M., Liang, L., Murakami-Ishibe, S., Allis, C.D., and Patel, D.J. (2007). Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. Mol Cell *28*, 677-691.

Liu, L., Oliveira, N.M.M., Cheney, K.M., Pade, C., Dreja, H., Bergin, A.M.H., Borgdorff, V., Beach, D.H., Bishop, C.L., Dittmar, M.T., *et al.* (2011). A whole genome screen for HIV restriction factors. Retrovirology *8*.

Liu, X., and Marmorstein, R. (2007). Structure of the retinoblastoma protein bound to adenovirus E1A reveals the molecular basis for viral oncoprotein inactivation of a tumor suppressor. Genes Dev *21*, 2711-2716.

Lomvardas, S., and Thanos, D. (2001). Nucleosome sliding via TBP DNA binding in vivo. Cell *106*, 685-696.

Lu, R., Yang, P., Padmakumar, S., and Misra, V. (1998). The herpesvirus transactivator VP16 mimics a human basic domain leucine zipper protein, luman, in its interaction with HCF. J Virol *72*, 6291-6297.

Lu, Y., Qian, X.Y., and Krug, R.M. (1994). The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. Genes Dev *8*, 1817-1828.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes. Curr Opin Genet Dev *13*, 127-135.

Luger, K., Mader, 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.

Lusic, M., Marcello, A., Cereseto, A., and Giacca, M. (2003). Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. Embo Journal *22*, 6550-6561.

Matzinger, P. (1994). Tolerance, danger, and the extended family. Annu Rev Immunol *12*, 991-1045.

Meng, X., Webb, P., Yang, Y.F., Shuen, M., Yousef, A.F., Baxter, J.D., Mymryk, J.S., and Walfish, P.G. (2005). E1A and a nuclear receptor corepressor splice variant (N-CoRI) are thyroid hormone receptor coactivators that bind in the corepressor mode. Proc Natl Acad Sci U S A *102*, 6267-6272.

Meunier, J.C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J., and Wychowski, C. (1999). Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex. J Gen Virol *80 (Pt 4)*, 887-896.

Miao, F., Li, S., Chavez, V., Lanting, L., and Natarajan, R. (2006). Coactivator-associated arginine methyltransferase-1 enhances nuclear factor-kappaB-mediated gene transcription through methylation of histone H3 at arginine 17. Mol Endocrinol *20*, 1562-1573.

Mohrmann, L., and Verrijzer, C.P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochim Biophys Acta *1681*, 59-73.

Morse, R.H. (1989). Nucleosomes inhibit both transcriptional initiation and elongation by RNA polymerase III in vitro. EMBO J *8*, 2343-2351.

Mueller, C.L., Porter, S.E., Hoffman, M.G., and Jaehning, J.A. (2004). The Paf1 complex has functions independent of actively transcribing RNA polymerase II. Molecular cell *14*, 447-456.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional Activators Enhance Polyadenylation of mRNA Precursors. Molecular cell *41*, 409-418.

Nakanishi, S., Lee, J.S., Gardner, K.E., Gardner, J.M., Takahashi, Y., Chandrasekharan, M.B., Sun, Z.W., Osley, M.A., Strahl, B.D., Jaspersen, S.L., *et al.* (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J Cell Biol *186*, 371-377.

Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C.W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., *et al.* (2010). Suppression of inflammation by a synthetic histone mimic. Nature *468*, 1119-1123.

Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., *et al.* (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol Cell *9*, 1201-1213.

Nordick, K., Hoffman, M.G., Betz, J.L., and Jaehning, J.A. (2008). Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II. Eukaryot Cell *7*, 1158-1167.

O'Connor, M.J., Zimmermann, H., Nielsen, S., Bernard, H.U., and Kouzarides, T. (1999). Characterization of an E1A-CBP interaction defines a novel transcriptional adapter motif (TRAM) in CBP/p300. J Virol *73*, 3574-3581.

Obenauer, J.C., Denson, J., Mehta, P.K., Su, X., Mukatira, S., Finkelstein, D.B., Xu, X., Wang, J., Ma, J., Fan, Y., *et al.* (2006). Large-scale sequence analysis of avian influenza isolates. Science *311*, 1576-1580.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Orphanides, G., and Reinberg, D. (2000). RNA polymerase II elongation through chromatin. Nature *407*, 471-475.

Palese P, S.M. (2007). Orthomyxoviridae: The Viruses and their Replication. In, H.P. Knipe DM, ed. (Philadelphia, Lippincott Williams & Wilkins), pp. 1647-1689.

Palese, P., and Wang, T.T. (2011). Why do influenza virus subtypes die out? A hypothesis. MBio *2*.

Panne, D., Maniatis, T., and Harrison, S.C. (2007). An atomic model of the interferon-beta enhanceosome. Cell *129*, 1111-1123.

Parekh, B.S., and Maniatis, T. (1999). Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. Molecular cell *3*, 125-129.

Pelka, P., Ablack, J.N., Fonseca, G.J., Yousef, A.F., and Mymryk, J.S. (2008). Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes. J Virol *82*, 7252-7263.

Penheiter, K.L., Washburn, T.M., Porter, S.E., Hoffman, M.G., and Jaehning, J.A. (2005). A Posttranscriptional Role for the Yeast Paf1-RNA Polymerase II Complex Is Revealed by Identification of Primary Targets. Molecular cell *20*, 213-223.

Pfennig, K.S. (2001). Evolution of pathogen virulence: the role of variation in host phenotype. Proc Biol Sci *268*, 755-760.

Pierson, T., McArthur, T., and Siliciano, R.F. (2000). Reservoirs for HIV-1: Mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. Annual Review of Immunology *18*, 665-708.

Piro, A.S., Mayekar, M.K., Warner, M.H., Davis, C.P., and Arndt, K.M. (2012). Small region of Rtf1 protein can substitute for complete Paf1 complex in facilitating global histone H2B ubiquitylation in yeast. Proc Natl Acad Sci U S A *109*, 10837-10842.

Pornillos, O., Alam, S.L., Davis, D.R., and Sundquist, W.I. (2002). Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. Nat Struct Biol *9*, 812-817.

Pruneski, J.A., Hainer, S.J., Petrov, K.O., and Martens, J.A. (2011). The Paf1 complex represses SER3 transcription in Saccharomyces cerevisiae by facilitating intergenic transcription-dependent nucleosome occupancy of the SER3 promoter. Eukaryot Cell *10*, 1283-1294.

Qiu, H., Hu, C., Gaur, N.A., and Hinnebusch, A.G. (2012). Pol II CTD kinases Bur1 and Kin28 promote Spt5 CTR-independent recruitment of Paf1 complex. EMBO J *31*, 3494-3505.

Qiu, H., Hu, C., Wong, C.M., and Hinnebusch, A.G. (2006). The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II. Mol Cell Biol *26*, 3135-3148.

Qiu, Y., Nemeroff, M., and Krug, R.M. (1995). The influenza virus NS1 protein binds to a specific region in human U6 snRNA and inhibits U6-U2 and U6-U4 snRNA interactions during splicing. RNA *1*, 304-316.

Quinlivan, M., Zamarin, D., Garcia-Sastre, A., Cullinane, A., Chambers, T., and Palese, P. (2005). Attenuation of equine influenza viruses through truncations of the NS1 protein. J Virol *79*, 8431-8439.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A Unifying Model for the Selective Regulation of Inducible Transcription by CpG Islands and Nucleosome Remodeling. Cell *138*, 114-128.

Ramirez-Carrozzi, V.R., Nazarian, A.A., Li, C.C., Gore, S.L., Sridharan, R., Imbalzano, A.N., and Smale, S.T. (2006). Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response. Genes Dev *20*, 282-296.

Rosenberg, B.R., Hamilton, C.E., Mwangi, M.M., Dewell, S., and Papavasiliou, F.N. (2011). Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. Nat Struct Mol Biol *18*, 230-236.

Rozenblatt-Rosen, O., Nagaike, T., Francis, J.M., Kaneko, S., Glatt, K.A., Hughes, C.M., LaFramboise, T., Manley, J.L., and Meyerson, M. (2009). The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3 ' mRNA processing factors. P Natl Acad Sci USA *106*, 755-760.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007a). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol *8*, 983-994.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007b). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Bio *8*, 983-994.

Saccani, S., Pantano, S., and Natoli, G. (2001). Two waves of nuclear factor kappaB recruitment to target promoters. J Exp Med *193*, 1351-1359.

Saccani, S., Pantano, S., and Natoli, G. (2002). p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment. Nat Immunol *3*, 69-75.

Sampath, S.C., Marazzi, I., Yap, K.L., Krutchinsky, A.N., Mecklenbrauker, I., Viale, A., Rudensky, E., Zhou, M.M., Chait, B.T., and Tarakhovsky, A. (2007). Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell *27*, 596-608.

Sarmento, L., Wasilenko, J., and Pantin-Jackwood, M. (2010). The Effects of NS Gene Exchange on the Pathogenicity of H5N1 HPAI Viruses in Ducks. Avian Dis *54*, 532-537.

Schaeper, U., Boyd, J.M., Verma, S., Uhlmann, E., Subramanian, T., and Chinnadurai, G. (1995). Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation. Proc Natl Acad Sci U S A *92*, 10467-10471.

Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. Nature *472*, 481-485.

Seong, S.Y., and Matzinger, P. (2004). Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. Nat Rev Immunol *4*, 469-478.

Shapira, S.D., Gat-Viks, I., Shum, B.O., Dricot, A., de Grace, M.M., Wu, L., Gupta, P.B., Hao, T., Silver, S.J., Root, D.E.*, et al.* (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell *139*, 1255-1267.

Sharp, G.W.G., and Hynie, S. (1971). Stimulation of Intestinal Adenyl Cyclase by Cholera Toxin. Nature *229*, 266-+.

Shelton, H., Smith, M., Hartgroves, L., Stilwell, P., Roberts, K., Johnson, B., and Barclay, W. (2012). An influenza reassortant with polymerase of pH1N1 and NS gene of H3N2 influenza A virus is attenuated in vivo. Journal of General Virology *93*, 998-1006.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006a). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006b). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Shogren-Knaak, M., and Peterson, C.L. (2006). Switching on chromatin - Mechanistic role of histone H4-K16 acetylation. Cell Cycle *5*, 1361-1365.

Simic, R., Lindstrom, D.L., Tran, H.G., Roinick, K.L., Costa, P.J., Johnson, A.D., Hartzog, G.A., and Arndt, K.M. (2003). Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. EMBO J *22*, 1846-1856.

Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem *280*, 41789-41792.

Sims, R.J., 3rd, Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Mol Cell *28*, 665-676.

Sixma, T.K., Pronk, S.E., Kalk, K.H., Wartna, E.S., Vanzanten, B.A.M., Witholt, B., and Hol, W.G.J. (1991). Crystal-Structure of a Cholera Toxin-Related Heat-Labile Enterotoxin from Escherichia-Coli. Nature *351*, 371-377.

Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., and Benkirane, M. (2010). HIV-1 Tat Assembles a Multifunctional Transcription Elongation Complex and Stably Associates with the 7SK snRNP. Molecular cell *38*, 439-451.

Squazzo, S.L., Costa, P.J., Lindstrom, D.L., Kumer, K.E., Simic, R., Jennings, J.L., Link, A.J., Arndt, K.M., and Hartzog, G.A. (2002). The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. EMBO J *21*, 1764-1774.

Stetson, D.B., and Medzhitov, R. (2006). Type I interferons in host defense. Immunity *25*, 373-381.

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

Suarez, D.L., and Perdue, M.L. (1998). Multiple alignment comparison of the non-structural genes of influenza A viruses. Virus Res *54*, 59-69.

Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev *19*, 815-826.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature Structural & Molecular Biology *14*, 1025-1040.

Thomas, J.O. (1999). Histone H1: location and role. Curr Opin Cell Biol *11*, 312-317.

Thomas, M., Kranjec, C., Nagasaka, K., Matlashewski, G., and Banks, L. (2011). Analysis of the PDZ binding specificities of Influenza A Virus NS1 proteins. Virol J *8*.

Tonna, S., El-Osta, A., Cooper, M.E., and Tikellis, C. (2010). Metabolic memory and diabetic nephropathy: potential role for epigenetic mechanisms. Nat Rev Nephrol *6*, 332-341.

Tran, H.G., Steger, D.J., Iyer, V.R., and Johnson, A.D. (2000). The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. EMBO J *19*, 2323-2331.

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

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tsai, B., Rodighiero, C., Lencer, W.I., and Rapoport, T. (2001). Protein disulfide isomerase (PDI) acts as a redox-dependent chaperone to unfold cholera toxin (CT). Gastroenterology *120*, A697-A697.

Tscherne, D.M., and Garcia-Sastre, A. (2011). Virulence determinants of pandemic influenza viruses. J Clin Invest *121*, 6-13.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. Bioessays *22*, 836-845.

Twu, K.Y., Kuo, R.L., Marklund, J., and Krug, R.M. (2007). The H5N1 influenza virus NS genes selected after 1998 enhance virus replication in mammalian cells. Journal of Virology *81*, 8112-8121.

Van Roey, K., Dinkel, H., Weatheritt, R.J., Gibson, T.J., and Davey, N.E. (2013). The switches.ELM resource: a compendium of conditional regulatory interaction interfaces. Sci Signal *6*, rs7.

Varble, A., Chua, M.A., Perez, J.T., Manicassamy, B., Garcia-Sastre, A., and tenOever, B.R. (2010). Engineered RNA viral synthesis of microRNAs. Proc Natl Acad Sci U S A *107*, 11519-11524.

Vignali, M., and Workman, J.L. (1998). Location and function of linker histones. Nat Struct Biol *5*, 1025-1028.

Wang, W., and Krug, R.M. (1998). U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the influenza virus NS1 protein. RNA *4*, 55-64.

Warner, M.H., Roinick, K.L., and Arndt, K.M. (2007). Rtf1 is a multifunctional component of the Paf1 complex that regulates gene expression by directing cotranscriptional histone modification. Mol Cell Biol *27*, 6103-6115.

Wasilenko, J.L., Lee, C.W., Sarmento, L., Spackman, E., Kapczynski, D.R., Suarez, D.L., and Pantin-Jackwood, M.J. (2008). NP, PB1, and PB2 viral genes contribute to altered replication of H5N1 avian influenza viruses in chickens. Journal of Virology *82*, 4544-4553.

Wild, G., Gardner, A., and West, S.A. (2009). Adaptation and the evolution of parasite virulence in a connected world. Nature *459*, 983-986.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. Journal of Biological Chemistry *278*, 34739-34742.

Wu, J.C., and Xu, W. (2012). Histone H3R17me2a mark recruits human RNA Polymerase-Associated Factor 1 Complex to activate transcription. P Natl Acad Sci USA *109*, 5675-5680.

Wysocka, J. (2006). Identifying novel proteins recognizing histone modifications using peptide pull-down assay. Methods *40*, 339-343.

Xhemalce, B., and Kouzarides, T. (2010). A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. Genes Dev *24*, 647-652.

Yang, Z.Y., Yik, J.H.N., Chen, R.C., He, N.H., Jang, M.K., 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.

Ye, H., Park, Y.C., Kreishman, M., Kieff, E., and Wu, H. (1999). The structural basis for the recognition of diverse receptor sequences by TRAF2. Mol Cell *4*, 321-330.

Yun, M.Y., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. Cell Research *21*, 564-578.

Zanier, K., Charbonnier, S., Sidi, A.O., McEwen, A.G., Ferrario, M.G., Poussin-Courmontagne, P., Cura, V., Brimer, N., Babah, K.O., Ansari, T.*, et al.* (2013). Structural

basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. Science *339*, 694-698.

Zhang, Y., Dasgupta, J., Ma, R.Z., Banks, L., Thomas, M., and Chen, X.S. (2007). Structures of a human papillomavirus (HPV) E6 polypeptide bound to MAGUK proteins: mechanisms of targeting tumor suppressors by a high-risk HPV oncoprotein. J Virol *81*, 3618-3626.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev *15*, 2343-2360.

Zhou, J.S., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nature Structural & Molecular Biology *14*, 1070-1076.

Zhou, K., Kuo, W.H., Fillingham, J., and Greenblatt, J.F. (2009). Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. Proc Natl Acad Sci U S A *106*, 6956-6961.