The “Histone Mimicry” by Pathogens

UWE SCHAEFER,1,4 JESSICA S.Y. HO,1,2,4 RAB K. PRINJHA,3 AND ALEXANDER TARAKHOVSKY1,5

1Laboratory of Immune Cell Epigenetics and Signaling, The Rockefeller University, New York, New York 10065
2Laboratory of Methyltransferases in Development and Disease, Institute of Molecular and Cell Biology (IMCB), Singapore 138673
3Epinova DPU, Immuno-Inflammation Therapy Area, GlaxoSmithKline, Medicines Research Centre, Stevenage SG1 2NY, United Kingdom
4Correspondence: tarakho@mail.rockefeller.edu

One of the defining characteristics of human and animal viruses is their ability to suppress host antiviral responses. Viruses express proteins that impair the detection of viral nucleic acids by host pattern-recognition receptors, block signaling pathways that lead to the synthesis of type I interferons and other cytokines, or prevent the activation of virus-induced genes. We have identified a novel mechanism of virus-mediated suppression of antiviral gene expression that relies on the presence of histone-like sequences (histone mimics) in viral proteins. We describe how viral histone mimics can interfere with key regulators of gene expression and contribute to the suppression of antiviral responses. We also describe how viral histone mimics can facilitate the identification of novel mechanisms of antiviral gene regulation and lead to the development of drugs that use histone mimicry for interference with gene expression during diseases.

Viral infection leads to a rapid onset of two major mutually antagonistic processes. Encounter of human or mouse cells with virus results in rapid activation of intracellular signaling networks that trigger expression of genes responsible for the elimination of the virus and the establishment of lasting antiviral immunity (Katze et al. 2002; Garcia-Sastre and Biron 2006; Kawai and Akira 2006). In turn, viruses can target the antiviral signaling events for the benefit of virus replication and long-term virus propagation (Katze et al. 2002; Garcia-Sastre and Biron 2006; Kawai and Akira 2006).

In the nucleus, virus–host antagonism evolves around key regulatory processes that govern gene transcription and DNA replication. The structural and nonstructural proteins of numerous viruses accumulate in the nuclei of infected cells, where they interact with distinct chromatin proteins, including histones, transcription factors, regulators of DNA repair, cell cycle transition, mRNA splicing, export, and catabolism (Nemeroff et al. 1998; Ballestas et al. 1999; Friborg et al. 1999; Fujimuro et al. 2003; Krug et al. 2003; You et al. 2004; Zhao and Elder 2005; Satterly et al. 2007; Ashour et al. 2009; Colpitts et al. 2011; Fonseca et al. 2012; Banerjee et al. 2013).

One of the most common antithost viral strategies involves the disruption of host protein–protein interactions by viral mimics of cellular proteins (Elde and Malik 2009; Davey et al. 2011; Tarakhovsky 2013). In particular, the inherent functions and properties of host signaling and regulatory networks make them extremely attractive targets for viral mimic-derived interference. For one, such networks typically rely on highly specific and dynamic protein–protein interactions to be able to quickly transmit information within the cell (Smock and Gierasch 2009; Stein et al. 2009). This is in part supported by the use of low-affinity protein–protein interactions that involve compact (3–20 amino acids long), degenerate and evolving modules that have been defined as short linear motifs (SLiMs) (Diella et al. 2008; Davey et al. 2012; Edwards et al. 2012; Van Roey et al. 2012, 2013; Weatheritt et al. 2012a,b). Most of the binding specificity and affinity of a SLiM is embedded within a 2–5-residue-long core of amino acids. This paucity of amino acids limits the strength of SLiM-mediated interactions to a relatively low affinity range (1–150 μM) that supports transiency and reversibility of SLiM-mediated protein–protein binding (Diella et al. 2008; Davey et al. 2012; Edwards et al. 2012; Van Roey et al. 2012, 2013; Weatheritt et al. 2012a,b; Dinkel et al. 2014).

The described features of SLiMs match well the features of the short protein-binding sequences within the amino-terminal domains of histone proteins (histone tail). The histone tail could be viewed as a collection of multiple overlapping SLiMs, with each motif (and its modification state) functioning as a discrete unit of information for histone-binding proteins (Fischle et al. 2003; Ruthenburg et al. 2007; Taverna et al. 2007). This feature of the histone tail may contribute to its unique capability to facilitate binding of numerous nonhistone proteins. However, the minimal size of the interacting histone sequences can potentially compromise the specificity. Therefore, it is tempting to speculate that histone-tail-dependent formation of the chromatin complexes relies on a kinetic proof-reading mechanism that has been initially introduced by Hopfield (1974) and Ninio (1975) as a mechanism for...
error correction in biochemical processes. This proofreading mechanism connects the small differences in binding energy of interacting partners and the amount of extra free energy required to generate the desirable reaction product. When undesirable reactants participate in a reaction, free energy consumption drives molecular circuits in cycles before completing the reaction (Hopfield 1974; Ninio 1975). It is plausible that the shortness of the histone tail and, even more, the minimal size of the functionally distinct histone tail subdomains, for example, H3K4 vs. H3K9 vs. H3K27, increases the length of time and amount of energy required for formation of functional transcription complexes. Although it is important to keep the error rates of transcription complex assembly low, the potential multiplicity of error-correcting cycles may increase the potentially damaging effect of “foreign” histone-like sequences that, by competing with cognate histone sequences, may prolong proofreading at the expense of cell fitness. In essence, although highly useful for the diversity of the chromatin complexes, the paucity of the amino acid sequence within the histone tail may create an opportunity for interference with the histone-based chromatin complexes by endogenous or exogenous histone-like SLiMs.

Here, we provide evidence for the existence and functional relevance of viral histone mimics. We also show that synthetic compounds that resemble histone mimics are able to interfere with the formation of transcription protein complexes and gene expression in cells of various types. Our studies suggest that targets of pathogen-derived or synthetic histone mimics can lead to identification of critical regulators of gene expression associated with inflammatory and viral diseases.

RESULTS AND DISCUSSION

Endogenous and Exogenous Histone Mimics

The existence of histone-like SLiMs (histone mimics) in nonhistone proteins has been initially shown in our earlier findings that reveal the presence of the histone H3-like sequence within the histone methyltransferase G9a that catalyzes dimethylation at lysine 9 of histone H3 (Sampath et al. 2007). G9a bears a 163-ARKT-166 motif that strongly resembles the 7-ARKS-10 motif of its target H3 target residue (Fig. 1A, left). Consistent with the presence of the histone mimic, G9a can automethylate itself on lysine 165 (Sampath et al. 2007), and this methylation facilitates the formation of the G9a bound repressor complex with chromodomain-containing protein HP1γ (Fig. 1A, right) (Sampath et al. 2007). The H3-like sequence in G9a is conserved in its homolog and heterodimerization partner GLP (Fig. 1A, left) (Tachibana et al. 2005; Sampath et al. 2007), although the two proteins share relatively poor primary sequence conservation in their amino-terminal domains. The initial discovery of histone mimics in G9a/GLP led to identification of numerous histone mimics in nuclear and nonnuclear proteins, where these mimics contribute to protein–protein interaction or protein stability (Lee et al. 2010, 2012; Donlin et al. 2012; Badeaux and Shi 2013; Shi et al. 2014).

The compactness and intrinsic evolutionary plasticity of SLiMs, including those that are present within histones, make them easy to imitate, thus enabling pathogen-derived SLiMs to compete with the host counterparts. The viral histone mimicry has been initially shown by identification of the histone H3-like sequence within the carboxy-terminal portion of the nonstructural protein 1 (NS1) derived from the H3N2 subtype of influenza A virus (Marazzi et al. 2012). We found that the carboxy-terminal domain of the NS1 protein carries the sequence 226-ARKS-229 (Fig. 1B, yellow) that resembles the first four amino acids of the histone H3 protein 1-ARTK-4 (red; Fig. 1B, top). Strikingly, both of these sequences were localized to the unstructured terminal domains of their respective proteins, indicating that the NS1 histone-like sequence is present within a similarly disordered structural context as the histone H3 tail.

The NS1 protein has a key role in virus propagation by suppressing the type I IFN response during influenza infection (Garcia-Sastre et al. 1998; Wang et al. 2000). Interestingly, among the 2753 unique full-length NS1 sequences obtained from the Influenza Virus Resource (Bao et al. 2008), the H3N2 subtype of viruses appears to be the predominant carrier of the bona fide histone H3K4-like sequence. In contrast, “tails” of NS1 proteins from other influenza subtypes vary significantly and display either no recognizable motifs or possess highly specific motifs, such as PDZ ligand (PL) motifs (Obenauer et al. 2006; Jackson et al. 2008; Thomas et al. 2011), or the sumoylation sequence within the NS1 tail of the H1N1 (1918) influenza A virus (Table 1) (Santos et al. 2013).

This NS1 tail diversity may contribute to unique features of the individual virus subtypes. The NS1 carboxy-terminal tail domain has previously been implicated as a virulence factor for some avian influenza A viral isolates (Jackson et al. 2008). The PL motifs ESEV and EPEV, found in the NS1 carboxyl terminus from multiple avian-derived influenza strains, have been associated with several highly pathogenic human isolates such as H5N1 1997 and H5N1 2003, respectively (Obenauer et al. 2006; Jackson et al. 2008; Thomas et al. 2011). These PL motifs enable viral interference with the apoptotic death of the infected cells, thus contributing to viral virulence (Golebiewski et al. 2011). Contrary to avian viruses, commonly circulating and relatively low-pathogenic human viruses predominantly contain PL motif sequences RSKV (H3N2 isolates) or RSEV (H1N1 isolates) that do not bind to known PDZ domain-containing proteins (Obenauer et al. 2006; Thomas et al. 2011). Furthermore, the overall loss of the NS1 carboxy-terminal tail from low-pathogenic viruses resulted in major attenuation of the virus (Jackson et al. 2008).

In the H3N2 strain, the presence of the histone mimic may contribute to the unique ability of the virus to compete with the cognate histone sequences for their common binding partners, including those that have an important role in regulation of antiviral gene expression. In support of this model, we found that both the histone H3 tail and the H3N2 NS1 tail bind in a sequence-dependent fashion to the Paf1 component of the polymerase-associated fac-
tor 1 complex (Paf1C) that contributes to RNA elongation as well as other cotranscriptional processes (Marazzi et al. 2012). As expected, the NS1 tails from influenza H5N1 (Marazzi et al. 2012) and H1N1 (Marazzi and Tarakhovsky, unpubl.) that do not carry the histone mimic did not bind to Paf1.

Overall, the NS1 tail could be viewed as a highly interactive internally disorganized motif, where amino acid variations contribute to target specificity and hence to the unique features of individual virus subtypes. One of the extreme extensions of this model is that the diversity of the entire NS1 tails collection of the global influenza pool matches the diversity of the histone tails as well as other “tail”-like disordered structures in other host proteins. This feature of NS1 may support the influenza subtypes’ ability to choose their optimal strategy.

![Figure 1](image.png)

**Figure 1.** The structure and function of histone mimics. (A) Methyltransferases G9a and GLP possess a functional histone mimic. The histone-like sequence (red letters) is localized within the amino-terminal domain of the proteins (Sampath et al. 2007). The G9a histone mimic methylation (red hexagons) is mediated in cis by the catalytic SET domain (Sampath et al. 2007) that is flanked by pre- and post-SET domains. The ankyrin repeat domain is involved in G9a interaction with methylated histones (Collins et al. 2008), and the methylated histone mimic in G9a (red hexagons) binds to the chromodomain-containing protein HP1γ that can also interact with methylated histone H3 (red hexagon). (B) The NS1 protein of the influenza A H3N2 virus possess a functional histone H3K4-like sequence. The histone-like sequence of NS1 (yellow letters) is localized within the nonstructured carboxyl terminus of the protein, whereas the homologous H3 sequence (red letters) is localized within the amino terminus of histone H3 (Marazzi et al. 2012). The NS1 histone mimic (yellow tails) is present in the nucleus (Greenspan et al. 1988), where it interacts with Paf1 and Chd1 proteins. Interaction with Chd1 depends on NS1 lysine methylation, whereas Paf1 can bind to the unmethylated or methylated NS1 histone mimic. The pattern of Paf1 and Chd1 binding to NS1 is similar to these protein interactions with histone H3. The schematic model describes a putative mechanism of NS1 interference with Paf1-mediated transcription of virus-induced genes.

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Year of appearance</th>
<th>NS1 carboxyl terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3N2</td>
<td>1968–2014</td>
<td>216 PKQKRK MART ARSKV</td>
</tr>
<tr>
<td>H1N1</td>
<td>1933–2014</td>
<td>216 PKQKRK MART IRSE V</td>
</tr>
<tr>
<td>H5N1</td>
<td>2003</td>
<td>216 PNQKRK MART IESE V</td>
</tr>
<tr>
<td>H1N1</td>
<td>1918 (pandemic)</td>
<td>216 PKQKRK MART IKSE V</td>
</tr>
<tr>
<td>H1N1</td>
<td>2009 (pandemic)</td>
<td>216 PKQK</td>
</tr>
</tbody>
</table>

The NS1 tails of the displayed influenza A subtypes vary at amino acid positions 217 and 226–229 (boxed). The bona fide H3K4-like sequence in NS1 of the influenza A H3N2 virus (226–229) is shown in yellow. The sumoylation sequence within the NS1 tail of the influenza A H1N1 1918 virus is shown in blue.
for the host-pathogen interaction, based on the genetic and epigenetic state of the infected host during the emergence of infection. In turn, the counterselection against certain viral subtypes, including some of the NS1 tail-less viruses, could be explained by the inability of NS1 to match the host factor(s), including those localized in the cell nucleus. Most provoking could be a consideration of the existence of the unique host-binding partner(s) for the sumoylated NS1 of the 1918 Spanish flu that was “removed” from the human population due to the exceptionally high morbidity of Spanish influenza.

Viral “Histone Mimics” Reveal Novel Regulators of Antiviral Response

NS1 binding to Paf1, although apparently specific for H3N2, may offer a clue for understanding the general mechanism that governs antiviral transcriptional responses. Pafl is an obligate member of the multiprotein Pafl complex that contributes to RNA elongation (Kim et al. 2010) in addition to other cotranscriptional processes. The latter include transcription-coupled histone modifications (Wood et al. 2003; Kim and Roeder 2009; Hahn et al. 2012) and regulation of mRNA 3’ end processing (Mueller et al. 2004; Penheiter et al. 2005; Nordrick et al. 2008; Rozenblatt-Rosen et al. 2009; Nagaike et al. 2011). In particular, the PaflC has been shown to impact mono-ubiquitylation of histone H2BK120 in human cells by promoting recruitment of the ubiquitin-conjugating enzyme Rad6 and the ubiquitin protein ligase Bre1 to chromatin (Wood et al. 2003; Kim and Roeder 2009). Rad6/Bre1-mediated monoubiquitylation of H2BK120 facilitates H3K4 trimethylation and H3K79 dimethylation by the Set1/MLL-1 and Dot1 complexes, respectively (Dover et al. 2002; Sun and Allis 2002; Kim and Roeder 2009; Nakanishi et al. 2009).

By binding to Pafl, the NS1 protein has the potential to interfere with multiple transcriptional processes, including those that support antiviral gene expression (Fig. 1B, bottom). Indeed, truncation of the Pafl-binding-NS1 sequence decreased virus ability to suppress antiviral gene expression and attenuated influenza infectivity in vitro (Marazzi et al. 2012). Conversely, the association with NS1 points to an important role of PaflC in antiviral transcriptional response.

In the course of cell infection with the influenza virus that lacks NS1 and hence does not interfere with Pafl function, Pafl is recruited de novo to ~200 genes associated with antiviral and antiviral response, including numerous type 1/III IFN-stimulated genes (ISGs) (Fig. 2A–C). The virus-induced recruitment of Pafl correlates with recruitment of RNA Pol II and an increase in Pafl-bound gene expression levels (Fig. 2C–E). Accordingly, the siRNA-mediated knockdown of Pafl suppressed selectively the expression of genes that recruit Pafl during infection but did not affect the expression of the majority of the ~10,000 genes that bind Pafl in a noninducible fashion (Fig. 2F).

The ability of the NS1 histone mimic to interact with host nuclear proteins could be expanded by its postranslational modification. Similar to histone H3, the lysine residue within the NS1 histone mimic could be acetylated or methylated in influenza-infected cells (Marazzi et al. 2012). Although this methylation appears to have no obvious impact on the NS1–Pafl interaction, the lysine methylation of NS1 supported its interaction with the chromatin-remodeling protein Chd1 (Marazzi et al. 2012). Moreover, the affinity of the methylated NS1 interaction with Chd1 is nearly equal to the affinity of Chd1 binding to the methylated lysine 4 of histone H3. Despite tight association of Chd1 with NS1, the siRNA-mediated deficiency of Chd1 had no effect on virus-induced gene expression. However, combined Pafl and Chd1 deficiencies resulted in a nearly complete abrogation of antiviral gene expression and 50-fold increase in influenza virus replication in vitro (data not shown). The major negative and cumulative effect of Chd1 and Pafl deficiencies on antiviral defense may reflect the Chd1 and PaflC interaction (Simic et al. 2003) that has been implicated in control of pre-mRNA splicing (Simic et al. 2007). NS1 interaction with Chd1 may provide the virus with access to the host splicing machinery and contribute to viral mRNA splicing and hence to viral replication (Chua et al. 2013).

PaflC and associated proteins contribute to replication of human adenovirus (HAdV). During HAdV infection, E1A recruits PaflC to the viral genome via Bre1, thus promoting viral early gene expression (Fonseca et al. 2013). The interaction of E1A with 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 the Pafl complex also facilitates Rad6/Bre1 activity, these data support the notion that PaflC and its related activities are key regulators of inflammatory response. Binding of the HIV protein TAT to PaflC has also been implicated in generating the permissive transcription environment for HIV replication (Sobhian et al. 2010). Reduced levels of HIV proviral integration in cells that overexpress Pafl highlighted the importance of Pafl expression in the HIV viral life cycle (Liu et al. 2011).

Functional Similarity between Synthetic and Natural Histone Mimes

The important role of PaflC in up-regulation of virus-inducible gene expression corroborates well with earlier data that showed the key role of transcriptional elongation in inflammatory gene expression. RNA elongation depends greatly on signal-induced lysine acetylation of histone H3 and H4 at gene promoters followed by the acetylated histone association with bromodomain and ET domain (BET) family of proteins (Hargreaves et al. 2009). The interaction between BET and acetylated histones is mediated by the evolutionary conserved ~110-amino-acid-long bromodomain that is present in each of the two tandem arranged modules (BD1 and BDII) in
Figure 2. Virus infection leads to Paf1 recruitment to the virus-activated gene loci. (A) A549 lung epithelia cells were infected with PR8/ΔNS1 virus that due to the lack of NS1 induces strong antiviral responses. The virus infection resulted in time-dependent increase (red) or decrease (blue) in Paf1 binding. The heatmap shows relative abundance of Paf1 binding to the gene transcriptional start sites (TSSs) in control or infected cells as determined by ChIP sequencing. Gray boxes at the right of the heat map indicate known interferon-stimulated genes (ISGs) (Schoggins et al. 2011). (B) The table shows the top five functional categories associated with genes that display >2-fold change in Paf1 binding at 12 h after infection. (C) Virus infection increases Paf1 abundance at key antiviral genes. Binding of H3K4me3 (green), Paf1 (blue), and Pol II (black) at virus-induced genes in noninfected cells (0 h) and PR8/ΔNS1-infected cells at 4 h and 12 h postinfection is shown. The y-axes represent the average number of tags per gene per 25 base pairs per 1,000,000 mapped reads. Scale values are indicated in parentheses. (D) Virus-induced Paf1 binding correlates with recruitment of RNA Pol II to antiviral gene loci. The plot shows the cumulative distribution function (CDF) of changes in Pol II binding to different groups of Pol II- or Paf1-bound genes at 12 h postinfection with influenza PR8/ΔNS1. Genes that change Paf1 binding upon infection: red line; all genes that bind Paf1: black line; all genes that bind Pol II: green line. The dashed gray line indicates a fold change of 1. (E) Virus-induced Paf1 binding correlates with up-regulation of antiviral gene expression. The plot shows the CDF of changes in gene expression for different groups of Paf1-associated genes at 12 h postinfection with influenza PR8/ΔNS1. Genes that change Paf1 binding upon infection: red line; all genes that bind Paf1: black line; all genes that do not bind to Paf1: blue line; all genes: yellow line. The dashed gray line indicates a fold change of 1. (F) Paf1 deficiency down-regulates selectively expression of genes that recruit Paf1 during infection. The plot shows the CDF of Paf1-binding levels (log 2 RPKM/bp) in noninfected A549 cells (left), in A549 cells 4 h postinfection with influenza PR8/ΔNS1 (middle), or 12 h postinfection (right) for genes that were suppressed by Paf1 knockdown (dark blue line), up-regulated by Paf1 deficiency (orange line), or for all genes (black line).
BRD4 as well as related BRDT, BRD2, and BRD3 proteins of the BET family (Dey et al. 2000, 2003; Kanno et al. 2004; Huang et al. 2007). The mechanism by which individual BETs contribute to gene regulation is not fully understood. BRD4 is the only BET protein that links acetylated chromatin to the P-TEFb complex that mediates the signal-induced phosphorylation and activation of RNA Pol II (Jang et al. 2005; Yang et al. 2005; Hargreaves et al. 2009; Zhou et al. 2012). BET proteins can also contribute to RNA synthesis by a bromodomain-independent association with Paf1C as well as with numerous effector proteins such as NSD3, a SET domain-containing histone methyltransferase; JMJD6, a histone arginine demethylase; or Chd4, a catalytic component of the NuRD nucleosome remodeling complex (Rahman et al. 2011; Liu et al. 2013).

We were the first to show the powerful anti-inflammatory effect of synthetic antagonists of BET bromodomains (Nicodeme et al. 2010). The pharmacological inhibitor of BET proteins that we named I-BET binds to the acetylated lysine-binding pockets of BET proteins and prevents these proteins from interacting with their cognate acetyl-lysine ligands (Nicodeme et al. 2010). In addition to suppression of inflammatory genes, the inhibitor of BET proteins (I-BET) possesses a remarkable capacity to inhibit expression of numerous inducible genes, including those that are driven by transcription factors such as c-Myc in tumor cells (Dawson et al. 2011; Delmore et al.

---

**Figure 3.** I-BET suppresses the antiviral response. (A) Venn diagrams display the number of virus-induced genes (>twofold, yellow circle) that were suppressed (>twofold, green circle) by I-BET treatment of A549 cells at 4 h or 12 h postinfection with influenza PR8/ΔNS1. The scatter plot shows gene expression levels in control (x-axis) vs. I-BET-treated cells (y-axis) at 4 h postinfection. The down- or up-regulated genes are shown in green or red, respectively. A black circle highlights genes that have an important role in antiviral response. The dashed black line indicates a twofold change in expression; the dotted black line indicates a 10-fold change in expression. (B) Venn diagrams display the number of IFNβ-induced genes (>twofold, orange circle) that were suppressed (>twofold, green circle) by I-BET treatment of A549 cells at 4 h or 12 h after incubation with IFNβ. The scatter plot shows gene expression levels in control (x-axis) vs. I-BET-treated cells (y-axis) at 4 h after IFNβ treatment. The down- or up-regulated genes are shown in green or red, respectively. A black circle highlights examples of genes that have an important role in the cellular antiviral response. The dashed black line indicates a twofold change in expression; the dotted black line indicates a 10-fold change in expression. (C) I-BET increases virus growth in vitro. The kinetics of influenza PR8/ΔNS1 virus replication in control (black circles) or I-BET-treated A549 cells (red circles) has been measured as described (Marazzi et al. 2012). The viral replication in I-BET-treated cells was compared with viral replication in Paf1-deficient A549 cells (blue triangles) that display significant increase in viral replication when compared with control cells (black triangles).
viral proteins that are homologous to the H3K4 or H3K9
polyprotein [Hepatitis C virus genotype 2] YP_001469630.1 ARSK 494 Carboxyl Varied
polyprotein [Hepatitis C virus genotype 5] YP_001469633.1 ARSK 494 Carboxyl Varied
HCV polyprotein [Hepatitis C virus genotype 4] YP_001469632.1 ARSK 494 Carboxyl Varied

Envelope glycoprotein B [Human herpesvirus 5] YP_081514.1 ARSK 257 Amino Membrane
NS5b [Hepatitis C virus genotype 3] YP_001396345.1 ARSK 4 Carboxyl Nuclear
Epitope glycoprotein B [Human herpesvirus 5] YP_001467685.1 ARSK 257 Amino Nuclear
HCV polyprotein [Hepatitis C virus genotype 4] YP_001396342.1 ARSK 494 Carboxyl Varied
Polyprotein [Hepatitis C virus genotype 5] YP_001396333.1 ARSK 494 Carboxyl Varied
Polyprotein [Hepatitis C virus genotype 2] YP_001396301.1 ARSK 494 Carboxyl Varied

Table 2. A variety of human pathogens bear putative histone-like sequences

<table>
<thead>
<tr>
<th>Viral protein with H3K4-like sequence</th>
<th>Accession</th>
<th>Motif</th>
<th>Distance from terminus</th>
<th>Terminus</th>
<th>Localization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonstructural protein 1 [Influenza A virus (A/New York/392/2004(H3N2))]</td>
<td>YP_308841.1</td>
<td>ARSK</td>
<td>4</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
<tr>
<td>DNA polymerase catalytic subunit [Human herpesvirus 6A]</td>
<td>NP_042931.1</td>
<td>ARSK</td>
<td>72</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>E2 protein [Semliki forest virus]</td>
<td>NP_0819006.1</td>
<td>ARSK</td>
<td>31</td>
<td>Carboxyl</td>
<td>Membrane</td>
</tr>
<tr>
<td>Attachment glycoprotein G [Human metapneumovirus]</td>
<td>YP_012612.1</td>
<td>ARSK</td>
<td>23</td>
<td>Amino</td>
<td>Membrane/cytoplasm</td>
</tr>
<tr>
<td>Truncated structural polyprotein [Semliki forest virus]</td>
<td>YP_006390078.1</td>
<td>ARSK</td>
<td>106</td>
<td>Carboxyl</td>
<td>Membrane/cytoplasm</td>
</tr>
<tr>
<td>NS5b [Hepatitis C virus genotype 3]</td>
<td>YP_001467685.1</td>
<td>ARSK</td>
<td>97</td>
<td>Amino</td>
<td>Perinuclear</td>
</tr>
<tr>
<td>Envelope glycoprotein B [Human herpesvirus 5]</td>
<td>YP_001467685.1</td>
<td>ARSK</td>
<td>257</td>
<td>Amino</td>
<td>Perinuclear</td>
</tr>
<tr>
<td>HCV polyprotein [Hepatitis C virus genotype 4]</td>
<td>YP_001467685.1</td>
<td>ARSK</td>
<td>494</td>
<td>Carboxyl</td>
<td>Varied</td>
</tr>
<tr>
<td>Polyprotein [Hepatitis C virus genotype 5]</td>
<td>YP_001467685.1</td>
<td>ARSK</td>
<td>494</td>
<td>Carboxyl</td>
<td>Varied</td>
</tr>
<tr>
<td>Polyprotein [Hepatitis C virus genotype 2]</td>
<td>YP_001467685.1</td>
<td>ARSK</td>
<td>494</td>
<td>Carboxyl</td>
<td>Varied</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Viral protein with H3K9/H3K27-like sequence</th>
<th>Accession</th>
<th>Motif</th>
<th>Distance from terminus</th>
<th>Terminus</th>
<th>Localization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol [Human adenovirus 35]</td>
<td>AP_000576.1</td>
<td>ARKS</td>
<td>28</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>DNA polymerase [Human adenovirus B]</td>
<td>AP_002213842.1</td>
<td>ARSK</td>
<td>28</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>Encapsidation protein IVA2 [Human adenovirus C]</td>
<td>NP_040515.1</td>
<td>ARKT</td>
<td>5</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
<tr>
<td>IVA2 [Human adenovirus 2]</td>
<td>AP_000165.1</td>
<td>ARKT</td>
<td>5</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
<tr>
<td>IVA2 [Human adenovirus 5]</td>
<td>AP_000201.1</td>
<td>ARKT</td>
<td>5</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
<tr>
<td>IV2 [Human adenovirus 1]</td>
<td>AP_000502.1</td>
<td>ARKT</td>
<td>5</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
<tr>
<td>E2 protein [Human papillomavirus type 88]</td>
<td>YP_001672011.1</td>
<td>ARKS</td>
<td>52</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>Single-stranded DNA-binding protein [Human adenovirus D]</td>
<td>YP_001974427.1</td>
<td>ARKT</td>
<td>72</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>ORF29 [Human herpesvirus 8]</td>
<td>YP_001129382.1</td>
<td>ARKT</td>
<td>126</td>
<td>Amino</td>
<td>Nuclear</td>
</tr>
<tr>
<td>BALF5 [Human herpesvirus 4]</td>
<td>YP_001712.1</td>
<td>ARKT</td>
<td>60</td>
<td>Carboxyl</td>
<td>Nuclear</td>
</tr>
</tbody>
</table>

The viral proteins that are homologous to the H3K4 or H3K9/H3K27 sequences in histone H3 are shown. The position of the putative viral histone mimics, with respect to the carboxyl or amino termini, as well as the cellular localization of the viral proteins, is indicated.

CONCLUSION

Our finding identifies the ability of viral and synthetic histone mimics to interfere with the formation of transcriptional complexes and gene expression. We show that the immunosuppressive NS1 protein of influenza A virus uses histone mimicry for suppression of antiviral gene expression. We also show the potent role of synthetic histone mimics in suppression of antiviral response.

The presence of histone mimics in structurally distinct viral proteins (Table 2) underscores the ability of viruses to use highly evolvable host protein sequences for developing the optimal strategy for pathogen–host interaction. By imitating histones, the viruses can challenge the adaptive capacity of the host that cannot modify the histone primary sequences without endangering the very foundation of eukaryotic cell organization. Therefore, it’s plausible that many of the histone-modifying enzymes that operate in the nucleus have a dual function, both as regulator of host gene expression as well as possible modifiers and attenuators of the viral histone mimic function.
In support of this model, acetylation of the histone mimic within NS1, that prevents its binding to Paf1C (Marazzi et al. 2012), could be seen as a part of the host defense against the virus. Accordingly, identification of enzymes that modify viral mimics may lead to the development of novel antiviral drugs.

Histone mimics may not only potentiate the ability of the virus to interfere with the host but also increase virus dependence on host transcription. It is tempting to speculate that by using histone mimicry the viruses may become more dependent on the host. As a consequence, the histone-mimic-bearing viruses or other pathogens may evolve eventually as symbiotic pathogens in a fashion similar to the commensal organisms in humans and animals.

The synthetic histone mimics such as the BET inhibitor I-BET or, similar to it, JQ1 (Filippakopoulos et al. 2010) can interfere with a significant number of genes largely overlapping but subtly different, including those triggered by viruses. This effect of BET inhibitors relies on a highly specific interaction with the bromodomains of the BET family of proteins. By using the analogy to SLiMs, it is possible to speculate that minor changes in the structure of BET inhibitors, although altering the inhibitor binding to BET proteins, may reveal therapeutically relevant drug targets that could not have been predicted based on the rational design that commonly ignores the founding role of SLiM in protein–protein interaction.

ACKNOWLEDGMENTS

We thank Adolfo Garcia-Sastre for providing the PR8/ANS1 virus; Raphael Cohn for performing ChIP-sequencing experiments; The Rockefeller University Genomics Resource Center; and Jonas Marcello and Ryan Kim for assistance with the data analysis. This work has been supported by a Lupus Foundation Grant (A.T.). The Rockefeller University, and the Epinova DPU, Immuno-sequencing experiments; The Rockefeller University Genesis, and the EpiCore DPU, Immuno-seq data analysis. The work has been supported by a Lupus Foundation Grant (A.T.). The Rockefeller University and the EpiCore DPU, Immuno-sequencing experiments.

REFERENCES


THE "HISTONE MICROMY" BY PATHOGENS


Jackson D, Hossain MJ, Hickman D, Perez DR, Lamb RA. 2008. The tumor suppressor CDC73 interacts with the P-TEFb component of P-TEFb and stimulates RNA polymerase II.


The "Histone Mimicry" by Pathogens

Uwe Schaefer, Jessica S.Y. Ho, Rab K. Prinjha, et al.

*Cold Spring Harb Symp Quant Biol* published online April 14, 2014
Access the most recent version at doi:10.1101/sqb.2013.78.020339

**License**

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](http://symposium.cshlp.org/subscriptions).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to: [http://symposium.cshlp.org/subscriptions](http://symposium.cshlp.org/subscriptions)

Copyright © 2013 Cold Spring Harbor Laboratory Press; all rights reserved