Viral enhancer mimicry of host innate-immune promoters

Citation for published version:

Digital Object Identifier (DOI):
10.1371/journal.ppat.1003804

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Plos pathogens

Publisher Rights Statement:
Copyright: © 2014 Kropp et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The inflammatory milieu is the natural habitat for a pathogenic infection, characterised by activity of pro-inflammatory signalling pathways and inflammatory cytokines. Viral entry rapidly activates a range of innate-immune signalling events such as the activation of Pattern Recognition Receptors (PRRs) [1–5]. A virus must therefore counteract intrinsic cellular and innate-immune responses to successfully complete the replication cycle. Frequently this is accomplished by encoding viral effector molecules that block these cellular responses by working as either structural or functional mimics of host target proteins [6–11]. Nuclear DNA viruses are dependent on the host transcriptional machinery to express the first viral genes; for example the immediate-early (IE) control elements of DNA viruses are by definition absolutely dependent on host transcription factors (TF) [12]. Therefore, these viruses are particularly hostage to their host transcriptional environment [13,14]. Here we propose that mimicry of regulatory DNA sequences by viral regulatory regions may also provide an additional strategy to counteract at IE times of infection the innate-immune response. In this context, viral IE control elements might functionally mimic innate-immune enhancers, taking advantage of the activated immune signalling TFs for promoting viral IE gene expression.

In other words: “If you can’t beat ‘em, Join ‘em.”

In exploring this possibility, we present a synopsis of the promoter-regulatory elements from seven extensively studied mammalian viruses with a DNA stage, and seven promoters representing prototypical cellular innate-immune genes. These are the SV-40 early enhancer, the E1A enhancer of HAdV5, the long terminal repeat (LTR) of HIV-1, the E6/7 long control region (LCR) of both HPV-16 and HPV-18, the major IE (MIE) enhancer of HCMV, and the enhancer-1 (Eh-1) regulatory region of HBV for viral sequences, and the enhancer regions of human IFNB1, IFNG, TNF, IRF1, IL8, IL12B, and IL1B for host sequences. First, we consider similarities between the primary sequence structures of the enhancers. Second, we present arguments for convergent evolution and structural flexibility inherent to enhancer sequences. Third, we discuss functional features and regulatory hallmarks that may be used to define viral enhancer mimicry of cellular immune enhancers.

Do Viral and Cellular Enhancers Display Any Primary Sequence Similarity?

To investigate if there is any similarity of primary sequences and therefore structural mimicry between the selected viral and cellular enhancers, we used the BLAST tool to compare the sequences against each other (Table 1) and applied an exhaustive pairwise multi-way alignment (CloneManager suite 7.0) to search for similarities in this group of sequences (Figure 1A). While multi-way alignment of the various selected viral and cellular promoter-regulatory regions (Figure 1A, top panel) reveals a lack of extended primary sequence homology, the pairwise BLAST comparison showed that small islands of sequence identity or high similarity are present (Table 1). We randomly compared some of these short sequence motifs with the JASPAR CORE (Vertebrates) database [15] and found that all checked motifs have similarities with consensus binding motifs for TFs (e.g., AP1, SP1, YY1, or RelA with relative scores of >0.8). This finding raises the question of whether there might be functional similarity. We therefore consider in the next section how convergent evolution of viral enhancers may have resulted in functional mimicry of the transcription control elements of innate-immune genes, providing a co-opting strategy for immune evasion.

Could Viral Regulatory Regions Evolve as Functional Mimics of Innate-Immune Enhancers without Extensive Sequence Similarity?

There are two principal genetic mechanisms that could lead to viral mimicry of host enhancers, horizontal transfer of cellular sequences to viral genomes or genetic drift of viral sequences. The first possibility, acquisition of cellular sequences through horizontal sequence transfer, could arise through illegitimate recombination with host DNA, for example by retro-transposition of non-coding RNA transcripts, resulting in the virus hijacking host transcription control sequences. If this were the general case, we would, however, expect significant structural similarity,
which we did not find in our analysis. Alternatively, but not mutually exclusive from horizontal transfer, viral enhancer mimics could arise through neutral evolution and genetic drift by sequence duplication or accumulation of point mutations. Duplicated sequence features are hallmarks for many viral and cellular enhancers [16–24]. For instance, deletion or loss of enhancer sequences in SV40 and JC polyomavirus promotes restoration of enhancer function through duplication of flanking sequences [25–28]. A third possibility is the accumulation of point mutations in enhancer sequences and subsequent fixation [29]. It has recently been described for a wide range of species that evolution of host-cell transcriptional control can occur in relatively short time spans and is mainly driven by the rapid and flexible emergence or loss of binding motifs rather than by evolution of the TF proteins themselves [30–36]. The described mechanisms of rapid enhancer evolution argue that viral enhancers could acquire functionality that mimics innate-immune enhancers without any extensive sequence homology, and this is consistent with the comparison of cellular and viral enhancers shown in Figure 1A. This possibility is underscored by the fact that promoter sequences seem to be poorly conserved even among members within a virus-family yet share many of the same regulatory elements [37]. For example the MIE enhancers of cytomegaloviruses show low levels of primary sequence similarity between the different species strains (Figure 1A, lower panel). Despite these differences, functionality of the enhancers is conserved between hosts for different CMV species strains, e.g., the human CMV enhancer can functionally complement deletion of the murine CMV enhancer [38] and human CMV enhancer sequences recapitulate in vivo biological sites of infection in species from mice to zebra fish [39–41].

What Features Would Classify a Viral Enhancer as an Innate-Immune Enhancer Mimic?

Since our work and that of others discussed so far indicates that viral enhancers are functional rather than structural mimics of host innate-immune enhancers, we suggest four principal hallmarks of functional enhancer mimicry. These are: 1) shared TF interactions independent of sequence structure, 2) similar kinetics of gene induction between cellular innate-immune and viral IE genes, 3) positive responsiveness to immune-stimulatory ligands, and 4) susceptibility to inhibition of

| Table 1. Summary of pairwise sequence comparison. |

<table>
<thead>
<tr>
<th>Virus</th>
<th>Selected enhancer region</th>
<th>Number of small islands of high similarity (BLAST)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCMV</td>
<td>Major immediate-early region</td>
<td>HCMV (2), IFNB1 (1), IL8 (1), IRF1 (1), TNF (1)</td>
</tr>
<tr>
<td>HPV-18</td>
<td>E6/7 Long control region</td>
<td>IFNG (1), IFNB1 (3), IL1B (2), IFN (1), IFNB1 (1)</td>
</tr>
<tr>
<td>HBV</td>
<td>Long terminal repeat region</td>
<td>IL12B (1), IL8 (1), IL1B (1), IFNB1 (1), IFNG (1)</td>
</tr>
<tr>
<td>HPV-16</td>
<td>E6/7 Long control region</td>
<td>IFNG (1), IL8 (2), IFNB1 (1), IL1B (1), TNF (1)</td>
</tr>
<tr>
<td>HPV-1</td>
<td>Long terminal repeat region</td>
<td>IFNG (1), IL8 (2), IFNB1 (1), IL1B (1), TNF (1)</td>
</tr>
<tr>
<td>SV-40</td>
<td>Major immediate-early region</td>
<td>IFNG (1), IFNB1 (2), IL8 (1), IRF1 (1), TNF (1)</td>
</tr>
</tbody>
</table>

Pairwise comparison of analysed enhancer sequences and number of similarity islands identified by BLAST (blastn) alignment. Each regulatory region was used as a reference sequence and compared to all other sequence elements. The number of sequence motifs with high similarity produced from this analysis are given in parentheses; comparisons that produced no significant similarities are not shown.

doi:10.1371/journal.ppat.1003804.t001
Table 2. List of identified interactions for the selected viral and host enhancers.

<table>
<thead>
<tr>
<th>TF Name</th>
<th>Entrez Gene ID</th>
<th>Protein Family</th>
<th>TF Name</th>
<th>Entrez Gene ID</th>
<th>Protein Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFKB1 (p50)</td>
<td>4790</td>
<td>NFkB</td>
<td>RelA (p65)</td>
<td>5970</td>
<td>NFkB</td>
</tr>
<tr>
<td>RelC</td>
<td>5966</td>
<td>NFkB</td>
<td>NFIB (p52)</td>
<td>4791</td>
<td>NFkB</td>
</tr>
<tr>
<td>C/EBP</td>
<td>N/A (generic)</td>
<td>C/EBP</td>
<td>CREB1</td>
<td>1385</td>
<td>bZIP</td>
</tr>
<tr>
<td>ATF1</td>
<td>466</td>
<td>AP</td>
<td>ATF2</td>
<td>1386</td>
<td>AP</td>
</tr>
<tr>
<td>ATF2</td>
<td>3725</td>
<td>AP</td>
<td>AP</td>
<td>2353</td>
<td>AP</td>
</tr>
<tr>
<td>AP1/Jun</td>
<td>3659</td>
<td>IRF</td>
<td>IRF1</td>
<td>6667</td>
<td>C2H2-zinc finger</td>
</tr>
<tr>
<td>SP1</td>
<td>6688</td>
<td>ETS</td>
<td>Sp1</td>
<td>7528</td>
<td>YY1</td>
</tr>
<tr>
<td>HMGI(Y)</td>
<td>3159</td>
<td>HMG</td>
<td>OCT 1</td>
<td>5451</td>
<td>OC/Pou</td>
</tr>
<tr>
<td>OCT 2</td>
<td>5452</td>
<td>OC/Pou</td>
<td>IRF2</td>
<td>3659</td>
<td>IRF</td>
</tr>
<tr>
<td>IRF3</td>
<td>3661</td>
<td>IRF</td>
<td>STAT1</td>
<td>6772</td>
<td>STAT</td>
</tr>
<tr>
<td>IRF7</td>
<td>3665</td>
<td>IRF</td>
<td>STAT2</td>
<td>6773</td>
<td>STAT</td>
</tr>
<tr>
<td>STAT3</td>
<td>6774</td>
<td>STAT</td>
<td>STAT4</td>
<td>6775</td>
<td>STAT</td>
</tr>
<tr>
<td>NFATp/NFATc</td>
<td>4773/511224</td>
<td>NFAT</td>
<td>YY1</td>
<td>7528</td>
<td>YY1</td>
</tr>
<tr>
<td>Y1</td>
<td>1051</td>
<td>bZIP</td>
<td>TBX21</td>
<td>30009</td>
<td>T-BOX</td>
</tr>
<tr>
<td>EOMES</td>
<td>8320</td>
<td>T-BOX</td>
<td>PPAR</td>
<td>N/A (generic)</td>
<td>Nuclear hormone receptor</td>
</tr>
<tr>
<td>PPAR</td>
<td>N/A (generic)</td>
<td>Nuclear hormone receptor</td>
<td>HNF1</td>
<td>6927</td>
<td>Hepatic nuclear factor</td>
</tr>
<tr>
<td>PPARG/PROX1</td>
<td>5468/5629</td>
<td>Nuclear hormone receptor</td>
<td>HNF3</td>
<td>2305</td>
<td>Hepatic nuclear factor</td>
</tr>
<tr>
<td>SMAD3</td>
<td>4088</td>
<td>SMAD</td>
<td>RUNX3</td>
<td>864</td>
<td>N/A</td>
</tr>
<tr>
<td>RUNX3</td>
<td>864</td>
<td>N/A</td>
<td>PRDM1/PRDI BF1</td>
<td>639</td>
<td>C2H2-zinc finger</td>
</tr>
<tr>
<td>HIVEP2/PRDI BF1</td>
<td>3097</td>
<td>C2H2-zinc finger</td>
<td>HIVEP1</td>
<td>3096</td>
<td>C2H2-zinc finger</td>
</tr>
<tr>
<td>NREBP</td>
<td>6651</td>
<td>N/A</td>
<td>NREBP</td>
<td>6651</td>
<td>N/A</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.ppat.1003804.t002
inflammatory signalling. In the following section we briefly discuss these hallmarks.

**Shared Transcription Factor Interactions**

The human genome encodes an estimated 1,700 to 1,900 TFs, with 1,391 representing high-confidence candidates [42]. These proteins represent an ample resource for viruses to harness. To probe, in more detail, the TF usage of the 14 viral and innate-immune enhancers selected (Table 1), we constructed unambiguous diagrams [43–45] of known TF interactions—available as an online resource on Figshare [46–52]. Using this approach we identified 72 interactions (Table 2) between the selected host and viral regulatory regions and host TFs. Of the 72 interactions identified, 43 were described for cellular enhancers and 50 for viral enhancers and 21 interactions (49% and 42% respectively) are shared among innate-immune enhancers and viral enhancers (Figure 1B). Annotation of this dataset using the BioMART tool (v0.7, ENSEMBL release v72) identified 31 TFs associated with “regulation of immune processes” (GO:0002376) in our 72 identified interactions. Notably, the extent to which the distinct viral enhancers share factors with the innate-immune genes varies (Figure 1C). This may be explained by the different physiological roles of the innate-immune genes and lifestyles of the selected viruses. Among the viruses, HCMV and HIV-1 enhancers show the largest TF overlap in total numbers of interactions with the innate-immune genes. In summary, we identified a substantial overlap in TF interactions between host and viral regulatory regions.

**Comparable Expression Kinetics**

It is noteworthy that host immediate-early response genes and viral immediate-early genes are, by definition, identified by the same criterion, namely that their expression is independent of newly synthesised proteins [12,53,54]. Upon infection of permissive cells, viral promoters are activated within the first hour of infection. This follows a typical expression profile for many innate immune genes in this dataset [59].

**Response to Immune-Stimulatory Ligands**

A corollary of viral enhancer mimicry of innate-immune regulatory functions is that the viral promoters/enhancers should be activated by the same signalling events as innate-immune genes. This implies that events during the infection process that trigger “antiviral” signalling cascades actually facilitate the initial viral transcription. In this context, it has been shown that activation of TLRs by LPS and CpG [60,61] increases activity in isolated HCMV-enhancer and HIV-LTR–driven reporter constructs [62–64]. This also seems to apply in the context of viral infection since cytokine signalling stimulates HBV gene expression [65] and HIV needs TLR-8 signalling in specific cell types for replication [66]. It is also notable that all of the viral control regions examined here have been shown to interact with AP-1 (Figure 1C). While AP-1 is not exclusively associated with innate-immune signalling, it can be activated by TLR signalling via MAPK-activation or by cAMP-related signalling during infection [67,68] and subsequently also binds to innate-immune enhancers. Taken together, these examples indicate that so called “antiviral” processes have the potential to facilitate viral IE gene expression and replication. In the future, their importance and potentially proviral role should be examined in viral infection models.

**Responsiveness to Negative Feedback Control**

Immune signalling pathways are tightly regulated by negative feedback with the inhibitors of signalling activity acting in a matter of minutes to hours [69,70]. Thus, innate-immune negative feedback loops should also inhibit viral gene expression and may play a role in viral latency. This hallmark of viral enhancer mimicry might prove the most challenging for scientific investigation. Interference with negative feedback regulators before infection may lead to an exacerbated immune response, either inducing an elevated antiviral state in the cell before the experimental infection or driving it into apoptosis. Still, proving that this hallmark is applicable to viral infections might provide new drug targets to inhibit viral infections. While, to our knowledge, no direct effects of negative regulators of inflammatory signalling on viral gene expression have been reported so far, it has been shown that anti-inflammatory drugs and chemical inhibitors of pro-inflammatory signalling, expected to increase viral replication, actually can inhibit viral gene expression and replication of HCMV, HBV, and HIV-1 [67,68,71–74].

**Concluding Remarks**

TFs activating innate-immune genes are regulated by PRR signalling that cannot be efficiently inhibited by viruses as their activation occurs during the viral entry process. Mimicking an innate-immune enhancer therefore has the advantage that TFs, already activated by the viral entry process, can be directly utilised in a time restricted manner to ensure viral gene expression at IE times. We hope this opinion opens debate and provides new insights for either reexamination or future-based investigations toward understanding viral gene activation and latency. Indeed we believe that the principle of viruses co-opting host-innate regulatory signals has broad implications toward understanding the biological role of viral enhancers, in acute and latent viral infections, and prospective host-directed antiviral therapeutic and vaccine strategies.

**Acknowledgments**

Our special thanks to Steven Watterson for organising the Figshare online resource, Kevin Robertson for editorial suggestions, Richard Perry for help with compiling the figures, and Venkatesh Mallikarjun, Chen Sz-Hau, Ming-Yuan Huang, David MacDonald, Aftab Yadav, Zuchra Zakirova, Julia Weber, and Mohammed Ab Abdullah for their student project reports (coordinated by Douglas Roy) on viral and cellular promoters as partial fulfilment of MSc by research in “Genomics and Pathway Biology.”

**References**

6. O’Neill LAJ, Golenbock D, Bowie AG (2013) The principle of viruses co-opting host-innate regulatory signals has broad implications toward understanding the biological role of viral enhancers, in acute and latent viral infections, and prospective host-directed antiviral therapeutic and vaccine strategies.
mimicry of native host ligands as a mechanism for host receptor engagement. Cell Host Microbe 14: 1–8.


