Published in final edited form as:

Nature. 2017 October 05; 550(7674): 46–47. doi:10.1038/nature24140.

Zapping viral RNAs

Stephen P. Goff

Departments of Biochemistry and Molecular Biophysics, and of Microbiology and Immunology, and at the Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA

Abstract

The discovery that the host defence protein ZAP specifically targets viral RNAs that are rich in a particular pair of adjacent bases — cytosine followed by guanine — sheds light on the evolution of viral RNA genomes.

The genomes of many RNA viruses contain a long-standing mystery: why does a particular dinucleotide, CG (a cytosine base followed by a guanine), occur at very low frequency^{1,2}? Vertebrate DNA genomes also have low CG levels, and the most common explanation is that the cytosine of CG is often modified by the addition of a methyl group — and the modified base is prone to mutation³.

But this CG methylation does not occur in RNA. The low CG content of viral RNAs also cannot be explained by any special constraints on the sequences that encode proteins, as might be supposed⁴. What could have driven RNA viruses to reduce the CG levels in their genomes? On page 124, Takata *et al.*⁵ propose an answer: the host defence protein ZAP (zinc-finger antiviral protein), which targets viral RNAs for destruction, specifically recognizes CG-rich sequences. The authors suggest that ZAP might be a major force of evolution that has led viruses to eliminate CG whenever possible.

ZAP was first identified in a screen for host proteins that restrict the expression of a mouse leukaemia-virus genome⁶. It binds viral RNAs using four zinc-finger domains and then recruits a molecular machine called the RNA exosome to mediate RNA degradation^{7,8}. ZAP restricts the replication of many viruses, including Ebola, hepatitis B and HIV-1. In each case, ZAP binds to specific regions of the viral RNAs, but its targets are not a simple sequence: the regions in each virus are large (about 400–500 nucleotides long) and, curiously, have no shared sequence motifs.

Takata *et al.* investigated the basis for the low CG content of RNA viruses in a study of HIV-1. The authors constructed 16 HIV-1 mutants in which different blocks of RNA were drastically changed to introduce as many mutations as possible without altering the amino-acid sequences of the encoded proteins. Some of the mutants, although viable, replicated poorly compared with wild-type viruses; viral RNA levels in infected cells were severely reduced. The authors noted that the sequence changes in these strains had created many CG pairs. They generated and analysed additional mutants that had high or low CG frequencies, and confirmed what has been shown^{1,2} in other viruses — that high CG levels in the virus strongly impair its replication.

To identify the host proteins that might be responsible for preventing replication, Takata and colleagues depleted host cells of several candidate proteins involved in RNA degradation. They found that depleting cells of ZAP protein or deleting the *ZAP* gene almost completely restored the mutant viruses to full replication competence. Might ZAP target CG-rich patches of RNA?

In what in my view is one of the prettiest experiments in their paper, the authors crosslinked ZAP to the viral RNA in infected cells, isolated the protein and sequenced the bound RNA. This revealed a near-perfect footprint of ZAP on CG-rich patches of mutant RNA, but not on the equivalent regions in controls. Small, shared sites of crosslinking mapped to rare CG-rich sites on both genomes. Next, the researchers demonstrated that embedding the CG-rich sequences in an unrelated gene rendered its expression sensitive to ZAP. These data show compellingly that ZAP binds and targets CG-rich sequences (Fig. 1).

Takata *et al.* boldly suggest that the basis for many viruses having evolved low-CG genomes is the selective pressure of ZAP. The fact that viral CG levels vary depending on which host a particular virus infects could now be laid at the feet of the ZAP activity in that host. In sum, the work reveals a key feature of the RNA elements that are recognized by ZAP, and offers an explanation for the long-standing mystery of why viral genomes often have low CG content.

These findings raise many interesting questions. For instance, is the protein interferon, which is often released by host cells in response to pathogen invasion, required for ZAP to impose selective pressure on viral RNAs? Basal ZAP levels are extremely low, and ZAP expression is induced by interferon; it might therefore be expected that the pressure exerted by ZAP would be imposed only on viruses that induce interferon expression in hosts. Curiously, interferon treatment does not increase the restriction of CG-rich picornaviruses², but perhaps these viruses induce interferon, and therefore ZAP, by themselves.

In this model, any virus that evaded detection, inactivated the interferon response or otherwise blocked ZAP would not be subject to ZAP's selective pressure. And, indeed, some viruses have taken this tack: the influenza NS1 protein inhibits ZAP activity^{9,10}, and the herpesvirus HSV-1 (a high-CG DNA virus) degrades ZAP messenger RNA (ref. 11). It remains unclear how other viruses with high CG content might evade ZAP.

Is ZAP the only factor that targets CG-rich sequences on viral RNAs? It seems to be the major force in the cells tested by Takata *et al.*, but there could be other factors in other cells, in other settings, at other times. For example, perhaps some sensors of viral RNAs, or the Dicer–Argonaute protein complexes that mediate RNA degradation by a different pathway, have a CG bias. There might be other factors that have yet to be discovered.

Takata and colleagues' work is likely to provoke new structural studies. Although a crystal structure of ZAP has been determined¹², co-crystals with RNA have not been obtained. It is likely that not all CGs are created equal in terms of being ZAP targets; clusters of CGs with special spacing and in special contexts are probably the best targets. Determining how CG-rich RNAs fit into the groove formed by ZAP's zinc fingers and make specific contacts with the protein will be of much interest.

The most far-reaching suggestion arising from this study is that ZAP constitutes a factor in the evolution of low CG content in the DNA of host cells. When CG-rich DNA sequences are transcribed, the resulting RNAs would be targeted by ZAP, preventing their function and so putting pressure on the genome to remove CG. ZAP would apply pressure only to transcribed regions of DNA, and so it is likely that other mechanisms, such as CG methylation, contribute as well. But once a low-CG state has been established by evolution, ZAP becomes free to attack non-self invaders that have high CG levels with minimal impact on the host. As is often the case, the study of viruses and virus–host warfare teaches us about the evolution of not only viral genomes, but also our own.

References

1. Atkinson NJ, Witteveldt J, Evans DJ, Simmonds P. Nucleic Acids Res. 2014; 42:4527–4545. [PubMed: 24470146]
2. Burns CC, et al. J. Virol. 2009; 83:9957–9969. [PubMed: 19605476]
3. Cooper DN, Gerber-Huber S. Cell Differ. 1985; 17:199–205. [PubMed: 3902251]
4. Tulloch F, Atkinson NJ, Evans DJ, Ryan MD, Simmonds P. eLife. 2014; 3:e04531. [PubMed: 25490153]
5. Takata MA, et al. Nature. 2017; 550:124–127. [PubMed: 28953888]
6. Gao G, Guo X, Goff SP. Science. 2002; 297:1703–1706. [PubMed: 12215647]
7. Guo X, Carroll JW, Macdonald MR, Goff SP, Gao G. J. Virol. 2004; 78:12781–12787. [PubMed: 15542630]
8. Guo X, Ma J, Sun J, Gao G. Proc. Natl Acad. Sci. USA. 2007; 104:151–156. [PubMed: 17185417]
9. Liu CH, Zhou L, Chen G, Krug RM. Proc. Natl Acad. Sci. USA. 2015; 112:14048–14053. [PubMed: 26504237]
10. Tang Q, Wang X, Gao G. J. Virol. 2017; 91:e01909–16. [PubMed: 27807230]
11. Su C, Zhang J, Zheng C. Virol. J. 2015; 12:203. [PubMed: 26625984]
12. Chen S, et al. Nature Struct. Mol. Biol. 2012; 19:430–435. [PubMed: 22407013]

Figure 1. ZAP protein selectively binds and blocks specific RNA viruses

a, Many RNA viruses have low levels of CG — a cytosine base followed by a guanine. Takata *et al.*⁵ demonstrate that this low CG content prevents the host-derived defence protein ZAP (zinc-finger antiviral protein) from binding viral RNA that enters the cell cytoplasm, and so enables viral replication. **b**, When the authors engineered an RNA virus to have high levels of CG, they found that ZAP selectively bound to CG-rich sequences through its four finger domains. Once bound, ZAP recruits a protein complex called the RNA exosome to degrade viral RNA, thus preventing viral replication.