Diverse roles and interactions of RNA structures during the replication of positive-stranded RNA viruses of humans and animals

Andrew Tuplin

School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK

Positive-stranded RNA viruses include important human, animal and plant pathogens. Their genomes are able to fold into complex structures stabilized by base pairing between individual nucleotides, many of which are highly conserved and have essential functions during virus replication. With new studies and technological advances the diversity of roles, mechanisms and interactions in which such structured viral RNA functions is becoming increasingly clear. It is also evident that many RNA structures do not function as discrete elements but through mechanisms involving multiple, long-range and often dynamic RNA–RNA interactions. Through a range of examples and recent advances, this review illustrates the diverse roles and mechanisms of structured viral RNA during the replication of positive-stranded RNA viruses infecting humans and animals.

INTRODUCTION

Positive-stranded RNA viruses are a phylogenetically diverse grouping including families such as the Flaviviridae, Coronaviridae andPicornaviridae. They possess single-stranded RNA genomes (~3–32 kb) that function as both genetic material and mRNA templates. With selection pressure to maximize their coding potential, their genomes have evolved a range of non-template functions involved in processes including virus translation, replication, subgenomic mRNA transcript production, encapsidation and modulation of host antiviral responses. Such non-template functions can be mediated through nucleotide sequence composition (Atkinson et al., 2014), specific motifs or more commonly via RNA structures interacting with host or viral trans-activating factors. RNA structures within virus genomes are stabilized by Watson/Crick base pairing and may comprise a simple stem loop with duplexed stem and single-stranded terminal-loop, or complex higher-order structures such as pseudoknots and dynamic long-range RNA–RNA interactions.

While their host range, structure and life cycle are extremely diverse, the requirement for positive-strand RNA virus genomes to act as both genetic material and mRNA remain in common, Fig. 1. Following release into the cytoplasm of a permissive cell, infecting genomes initially template translation of viral proteins including RNA-dependent RNA polymerase (RdRp), which catalyses synthesis of RNA negative-strand intermediates that in turn template synthesis of positive-strand RNA molecules. Daughter positive-strand RNA is multifunctional and templates further rounds of genome replication and translation or can be packaged into new virion particles. In many instances mechanisms for temporal and spatial regulation of these key events depend on structured RNA. Such functional RNA elements act through an extremely diverse range of mechanisms involving interaction with host/viral trans-activating factors and complex, often dynamic, RNA–RNA interactions. Our understanding of how such mechanisms function has until recently been limited, however with new structural mapping methods such as SHAPE (Merino et al., 2005; Wilkinson et al., 2006) and powerful in silico prediction algorithms, we are starting to develop a clearer picture of the interactions involved. Structured RNA and its dynamic interactions play an essential role in the replication cycle of positive-stranded RNA viruses infecting the full range of prokaryotic and eukaryotic hosts. However, due to their large number and the range of mechanisms in which they are involved, this review cannot be an exhaustive discussion of all functional RNA elements (for reviews of structured RNA in positive-strand RNA viruses of plants and prokaryotes, see Pathak et al. (2011), and Harvey et al. (2013) and Olsthoorn (2011), respectively). Consequently, using a range of examples and recent advances, this review illustrates the mechanistic and structural diversity of viral RNA elements involved in the replication of positive-strand RNA viruses that infect animal and human hosts.

TRANSLATION

Positive-strand RNA viruses compete with endogenous cellular mRNAs for host translational machinery. In many instances, these viruses have evolved functional RNA structures that impart a competitive advantage in recruiting
translation initiation factors. Alternatively, such structures may provide temporal control to translation events by modulating the efficiency of initiation or reprogramming translational extensions.

**Initiation**

Cellular translation is generally initiated by recognition of an m7G cap by eukaryotic initiation factor 4F (eIF4F) – a complex of eIF4A, eIF4E and eIF4G – that recruits the 40S ribosome preinitiation complex to the mRNA. Many viruses inhibit host cell translation through cleavage or modification of cap-dependent factors and may directly recruit canonical and non-canonical factors in a cap-independent process, via structured RNA designated internal ribosome entry sites (IRES). Based on structure, mechanism and initiation factor requirements, IRES elements can be grouped into four categories. Type I elements include poliovirus, type II foot and mouth disease virus (FMDV), type III hepatitis C virus (HCV) and type IV the intragenic region (IGR) cricket paralysis virus (CrPV) IRES. The 40S preinitiation complex either binds at an
upstream position and scans downstream to the authentic AUG start codon (type I) or interacts directly with the start codon (types II–IV). Depending on IRES type, 40S subunit binding is initiated through different combinations of canonical/non-canonical translation factors. However, type IV elements such as the CrPV IGR–IRES, interact directly with the 40S subunit, independent of cellular initiation factors (Jan & Sarnow, 2002).

Modulation

IRES elements are generally located within 5′ non-coding regions (NCRs), however their initiation function may be modulated by RNA structures or long-range interactions within other regions of the virus genome. FMDV translation is upregulated by a long-range interaction between its IRES and a region of the 3′NCR containing RNA structures SL1, SL2 and a poly(A) tract, Fig. 2(a) (López de Quinto et al., 2002; Serrano et al., 2006). Rather than specific stem loops, enhancement is dependent on the overall higher-order RNA folding structure adopted by the entire IRES and 3′NCR (Serrano et al., 2006). Interestingly, this structure is competitive with an alternative interaction between SL1–SL2 and the S-fragment stem loop in the 5′NCR. It has been postulated that the S-fragment is involved in genome replication and that the alternative mutually exclusive long-range interactions represent a mechanism modulating FMDV translation and replication (Serrano et al., 2006).

Similarly, regulation of HCV translation involves various complex (and potentially competing) higher-order RNA interactions, Fig. 2(b). Stem loop SL9266 [alternatively 5BSL3.2 (You et al., 2004) or SLV (Lee et al., 2004)] in the RdRp encoding region of the open reading frame (ORF), is the core of a dynamic pseudoknot (SL9266/PK), undergoing multiple RNA–RNA interactions involved in regulation of virus translation and replication. The subterminal bulge of SL9266 forms RNA–RNA interactions with both the IIId loop of the IRES (Romero-López & Berzal-Herranz, 2012) and an upstream region of the RdRp encoding region centred on nucleotide 9110 (Diviney et al., 2008). The two potentially competing interactions are mutually exclusive, with the IIId loop interaction inhibiting HCV translation (Romero-López & Berzal-Herranz, 2012) and that with the 9110 region essential for virus replication (Diviney et al., 2008). The terminal-loop of SL9266 forms a ‘kissing-loop’ interaction with stem loop SL9571 (alternatively SL2) in the 3′NCR (Friebe et al., 2005). The ‘kissing-loop’ is a dynamic structure forming open and closed conformations, both of which are essential for efficient virus replication (Tuplin et al., 2012). We have shown that the closed conformation stimulates HCV IRES initiated translation (Tuplin and others, unpublished results) and the open structure preferentially binds cellular protein EWSR1 upregulating genome replication (Oakland et al., 2013). It has been suggested that SL9266/PK functions in the temporal control of early translation and replication events (Diviney et al., 2008; Oakland et al., 2013). However, much remains to be understood regarding this complex mechanism of dynamic RNA–RNA and RNA–protein interactions, in particular how alternative conformations are stabilized and interact with different trans-activating factors.

Many positive-strand RNA viruses that use capped, rather than IRES initiated translation, also contain RNA structures that confer regulatory control and a competitive advantage over cellular mRNA initiation. The 3′NCR of flaviviruses such as dengue virus (DENV) contain two dumbbell-like stem loops (DB1 and DB2) that form local pseudoknots (DB/PK) necessary for optimal translation, Fig. 2(c) (Manzano et al., 2011). Eukaryotic m7G cap-dependent translation is regulated by interaction between the mRNA 3′ poly(A) tract and poly(A) binding protein (PABP) that recruits eIF4G, triggering initiation complex formation and cyclization. DENV lacks a poly(A) tract, however the DB/PKs independently bind PABP, enabling efficient initiation; even under conditions of reduced eIF4G (Polacek et al., 2009).

Reprogramming

Many positive-strand RNA viruses increase the coding capacity of their genome through translation reprogramming mechanisms such as stop codon read-through or programmed -1 ribosomal frameshifting. Reprogramming enables translation of C-terminal extended polypeptides and control over levels of translation products. Alphavirus non-structural proteins are translated as a polyprotein from a single ORF but in many instances (for example Sindbis virus) an in-frame UGA stop codon is located upstream of the RdRp coding region (Strauss et al., 1983). In a proportion of ribosomes, the stop signal is read-through in a mechanism stimulated by (but not dependent on) a large 3′ RNA structural element (Firth et al., 2011). Alternatively, ribosomal frameshifting generally depends on signals incorporating structured RNA and ‘slippery’ nucleotide motifs. Coronavirus (CoV) contain a signal in ORF1a that enables a proportion of elongating ribosomes to switch into the −1 reading frame (ORF1b), bypassing a stop codon and translating C-terminal extended ORF1a/b products (Giedroc & Cornish, 2009; Plant & Dinman, 2008). In SARS-CoV, the signal incorporates a pseudoknot containing three internal stems at which a proportion of ribosomes pause and stutter, shifting into the −1 reading frame at an upstream ‘slippery’ motif.

GENOME REPLICATION

Temporal and spatial regulation of negative-strand synthesis is frequently provided by structured RNA elements in the positive-strand 3′NCR, whilst growing evidence suggests that RNA structures in the negative-strand 3′NCR influence positive-strand synthesis. However, it is becoming increasingly clear that regulation also often involves RNA interactions and structure in both the 5′NCR and ORF. Such mechanisms can function via direct RNA–RNA
Fig. 2. RNA secondary structures and higher-order interactions within positive-strand RNA virus genomes. Simplified cartoon representations of viral genomes showing functional local RNA structures and higher-order RNA–RNA interactions between genome regions as dashed red lines. Breaks in genomes and regions not shown are represented by diagonal lines (-/-) and black-dashed lines, respectively. Relative genome positions are indicated below each RNA molecule by schematic genome maps showing NCRs and the ORF. (a) FMDV showing ‘kissing-loop’ interactions between the terminal loops of SL1 and SL2 and alternative long-distance interactions between different regions of the 3’NCR and either the S-fragment or IRES. Pk represents a region of predicted pseudoknots. (b) HCV showing a ‘kissing-loop’ interaction between the terminal loops of SL9266 and SL9571 and alternative long-range interactions between the subterminal bulge of SL9266 and either the IRES or upstream to a position centred on nucleotide position 9110. (c) DENV showing alternative linear and circular conformations and DB pseudoknot interactions. Unbroken red, yellow and blue lines represent complementary UAR (upstream AUG region), DAR
complementarity or RNA-binding protein intermediates and frequently initiate genome cyclization between the 5’ and 3’ NCRs.

**Flaviviruses**

Flaviviruses possess structurally dynamic genomes that switch between linear and circular conformations, both of which are essential for virus replication. Genome cyclization is stabilized by direct RNA–RNA interactions between inverted complementary repeats (UAR, DAR and CS motifs) at either end of the genome, Fig. 2(c) (Gebhard et al., 2011; Villordo & Gamarnik, 2009). In DENV, cyclization is essential for negative-strand synthesis (Alvarez et al., 2005a; Khromykh et al., 2001; Lo et al., 2003), and a balance between both conformations is critical for virus replication (Villordo et al., 2010). DENV genome replication is dependent on activation of viral RdRp via binding 5’NCR stem loop SLA and genome cyclization to reposition the activated replication complex at the 3’NCR transcription initiation site (Filomatori et al., 2011; Filomatori et al., 2006). Although SLA and genome cyclization are essential for DENV negative-strand synthesis they are not sufficient. Evidence from various laboratories indicates the involvement of further RNA rearrangements and structures in both NCRs and the ORF. For example, the 3’NCR DB/PKs enhance transcription and the base of stem loop 3’SL represses it due to overlapping with the 3’UAR cyclization signal (Filomatori et al., 2011; Manzano et al., 2011); consequently, the base of 3’SL must open to allow 5’-3’UAR binding and subsequent negative-strand synthesis.

Flavivirus cyclization is dependent on direct RNA–RNA complementarity rather than protein factors (Alvarez et al., 2005b). However, various cellular and viral proteins, essential to efficient virus replication, bind in structured regions of the genome and have been postulated to control or stabilize conformational rearrangement. In arboviruses such as DENV and alalphaviruses, stem loop mutants have been observed with different phenotypes in mammalian and mosquito cells (Fayzulin & Frolov, 2004; Gorchakov et al., 2004; Groat-Carmona et al., 2012; Villordo & Gamarnik, 2013). Such differences most likely reflect differences in trans-activating protein factor availabilities, although differences in intracellular environments such as temperature and cation concentrations may also play a role.

**Picornaviruses**

The genome of poliovirus contains various functional RNA structures such as cis replicating elements (cre) responsible for uridylylation of the small protein-nucleotide primer of replication, VPg. The genome also includes a highly structured 5’NCR containing a 5’ cloverleaf (5’CL) and IRES and two stem loops in the 3’NCR (SL-x and SL-y) that form a ‘kissing-loop’, Fig. 2(d) (Ogram & Flanagan, 2011). Cellular poly(rC) binding protein 2 (PCBP2) interacts with the poliovirus IRES during translation initiation and with stem loop 5’CL, as part of a ribonucleoprotein complex with viral protein 3CDpro (5’CL-RNP), during genome replication. Formation of the 5’CL-RNP is essential for initiation of negative-strand synthesis, with genome cyclization occurring through interaction between PCBP2 in the 5’CL-RNP and PABP bound to the 3’NCR. PCBP2 is cleaved by poliovirus proteinase 3CD rendering it unable to bind the IRES and preventing translation initiation, while binding to 5’CL and thus replication is unaffected. Indeed, cleavage of PCBP2 by 3CD is necessary for poliovirus replication and may constitute a switching mechanism between translation and RNA replication (Chase et al., 2014). The 3’NCR ‘kissing-interaction’ is also required for efficient replication and surprisingly its disruption has a much greater inhibitory effect than deleting the entire 3’NCR, possibly due to faulty positioning of the 5’CL-RNP in ‘kissing-loop’ mutants blocking successful initiation. Structures within the poliovirus genome are relatively well documented. However, the dynamics of their interactions within host cells and the role of other regions of structured RNA remain to be fully characterized. For example, two recent studies revealed a number of novel structures within the ORF (Burrill et al., 2013; Song et al., 2012), one of which – element 3D-7000 – is implicated in RNA replication through an as yet unknown mechanism (Burrill et al., 2013).

**EVASION OF INNATE IMMUNE RESPONSES**

Antiviral innate responses generally depend on recognition of viral RNA through pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs). Recognition of viral RNA PAMPs (such as double-stranded molecules, length, sequence, structure and location), trigger a range of antiviral responses. Consequently, viruses have evolved numerous countermeasures, often involving structured RNA, to block recognition by PRRs or overcome innate responses. For example, an RNA structure in the 3CDpro coding region of poliovirus (ciRNA) binds RNase L, competitively inhibiting its antiviral activity (Han et al., 2007; Keel et al., 2012). Many positive-stranded RNA viruses, such as HCV, exhibit extensive RNA structure designated genome-scale ordered RNA structure (GORS) (Simmonds et al., 2004). Distinct from discrete structural elements, GORS inhibits detection by cellular PRRs with its presence correlating to persistent viruses (Witteveldt et al., 2014). A number of PRR interferon induced proteins bind 5’ caps lacking 2’-O methylation, in alphaviruses secondary structures in their 5’NCR alter this binding and block their
antiviral effect (Hyde et al., 2014). Alternatively, highly structured subgenomic flavivirus RNAs (sfRNA) – produced during flavivirus infection as the product of host Xrn1 endonuclease activity terminating at a pseudoknot resistant to Xrn1 helicase activity (Chapman et al., 2014) – increase viral pathogenicity in mice, play a role in inhibition of a type I interferon responses and are involved in silencing RNAi in insect cells (Schnettler et al., 2012). Such varied mechanisms by which structured RNA interacts with or manipulates antiviral responses, illustrates its key role in the evolutionary arms race between viruses and host innate immunity.

**SUMMARY AND PERSPECTIVES**

Using a range of examples and recent advances, this review illustrates the diversity of mechanisms by which structured RNA influences different stages of positive-RNA virus replication. As our understanding of these mechanisms develops, in part through new methodologies enabling structural mapping of longer RNA molecules at higher resolutions, the importance and diversity of – often long-range and dynamic – higher-order RNA structures is becoming increasingly clear. For example, various studies have clearly demonstrated that flavivirus genomes switch between alternative RNA conformations and that a balance between the two is critical to virus replication. However, our understanding of the mechanisms and trans-activating factors involved in controlling or sensing such complex and often dynamic interactions is limited. Indeed, elucidation of transient RNA–RNA and RNA–protein interactions remains challenging and we have little comprehension of how essential RNA structural changes and dynamic interactions are regulated within different intracellular environments and compartments. In summary, recent advances have made a significant impact on our understanding of viral RNA structure and its diverse roles across a range of mechanisms essential to positive-strand RNA virus replication. With future studies refining and extending our understanding, it can be expected that the potential of structured RNA as an antiviral target will be further investigated and that a clearer understanding of their dynamics during virus replication will inform the study of RNA structures in other systems, such as eukaryotic mRNA and long non-coding RNA function.

**REFERENCES**


