PORTLAND PRESS

Review Article

Structure and regulation of coronavirus genomes: state-of-the-art and novel insights from SARS-CoV-2 studies

Ilaria Manfredonia and Danny Incarnato

Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

Correspondence: Danny Incarnato (d.incarnato@rug.nl)

OPEN ACCESS

Coronaviruses (CoV) are positive-sense single-stranded RNA viruses, harboring the largest viral RNA genomes known to date. Apart from the primary sequence encoding for all the viral proteins needed for the generation of new viral particles, certain regions of CoV genomes are known to fold into stable structures, controlling several aspects of CoV life cycle, from the regulation of the discontinuous transcription of subgenomic mRNAs, to the packaging of the genome into new virions. Here we review the current knowledge on CoV RNA structures, discussing it in light of the most recent discoveries made possible by analyses of the SARS-CoV-2 genome.

Introduction

RNA viruses represent one of the most common classes of pathogens behind human diseases [1], with around 180 currently recognized species, and around three new species discovered every year [2]. The harmfulness of these viruses is partially supported by their ability to rapidly evolve and adapt, allowing easier escape of host immune responses and quicker development of resistance towards drugs and vaccines. This ability relies on the low fidelity of viral RNA polymerases. The lack of any proofreading activity results in mutation rates as high as 10^{-3} [3], nearly 6–7 orders of magnitude higher than those of bacterial DNA polymerases [4].

The RNA genomes of these viruses carry two layers of information [5]. On one layer, the primary sequence encodes for all the viral proteins needed to hijack the host cell machinery and to generate new viral particles.

The second layer of information relies on the ability of RNA to fold into stable secondary structures, by base-pairing between regions of internal complementarity, followed by further compaction into tertiary structures. RNA structure elements, as well as their ability to dynamically interconvert between alternative conformations, have been widely proven to be essential players in the life cycle of various RNA viruses. As an example, the HIV trans-activation response (TAR) element, that can switch between two alternative conformations, is essential for the Tat-mediated activation of viral replication, it provides the docking site for the interaction with several host proteins, and it can be further processed into a mature microRNA to repress target host mRNAs [6–9].

In this perspective, recent technological advances based on RNA structure probing approaches coupled to next-generation sequencing (NGS) technologies [5,10,11] are progressively expanding our ability to query higher-order structures in the genome of RNA viruses, as well as to dissect their dynamics *in vivo* [12].

In this review, we will focus our attention on *Coronavirinae*, a subfamily of RNA viruses consisting of four genera, namely *Alphacoronaviruses* (alpha-CoV), *Betacoronaviruses* (beta-CoV), *Gammacoronaviruses* (gamma-CoV), and *Deltacoronaviruses* (delta-CoV) [13]. Alpha-CoV and beta-CoV are able to infect humans, usually resulting in respiratory illness [14]. Beta-CoV comprise the three most pathogenic

341

coronaviruses known to date: the severe acute respiratory syndrome virus (SARS-CoV), the Middle East respiratory syndrome virus (MERS-CoV), and the SARS-CoV-2 virus, responsible for the currently ongoing (December 2020) COVID-19 pandemic [15,16]. Coronaviruses (CoV) are characterized by the largest known positive-sense single-stranded RNA genomes (∼26–32 kb) [13], with a highly conserved architecture [17]. Despite the relatively poor fidelity of viral RNA-dependent RNA polymerases (RdRp), overall long-term integrity of such large genomes is ensured by the presence of the 3′–5′ exonuclease ExoN domain of the non-structural protein nsp14, that enables proof-reading at genome replication [18,19].

The genomic RNA (gRNA) has a 5′ cap and 3′ polyA tail, allowing direct translation of the non-structural polyproteins (nsp) ORF1a, and ORF1b, followed by assembly of the replicase-transcriptase complex (RTC). The RTC drives both genome replication and discontinuous transcription of subgenomic mRNAs (sgRNAs). Discontinuous transcription is mediated by short AU-rich transcription regulating sequences (TRSs), located either right downstream of the 5′ leader (TRS-L), or right upstream of each viral ORF (TRS-B), except for ORF1a and ORF1b [20]. The resulting sgRNAs are further translated to produce the structural proteins spike (S), envelope (E), membrane (M) and nucleocapsid (N), as well as several accessory proteins.

Over the past few months, enormous efforts have been put in trying to better understand the biology of the SARS-CoV-2 virus, to find an Achilles' heel and to confine the COVID-19 pandemic. In a very short time, the structure and function of the SARS-CoV-2 proteome [21–25], transcriptome [26], as well as the interactions with the host cell proteome, both at the protein [27–29] and RNA levels [30,31], have been dissected. We will here summarize the current knowledge on coronavirus RNA structures and their regulation, as well as the most recent discoveries, as revealed by cutting-edge high-throughput analyses conducted on SARS-CoV-2.

5′ UTR structures

The 5′ UTR of CoV is highly structured. A first consensus model of the 5′ UTR was originally proposed by comparative analysis of nine CoV from the three major CoV groups (alpha, beta, and gamma-CoV), and identified the three major conserved stem–loops SL1, SL2, and SL4 [32]. All three SLs have been consistently predicted and experimentally verified in SARS-CoV-2 [33–39]. Besides these, up to eight SLs can be present in CoV 5′ UTRs, with some degree of variation across different genera (Figure 1).

The SL1 has a functionally bipartite structure. On the one hand, the upper part of the SL1 must be folded. Indeed, destabilizing mutations introduced in the SL1 of Mouse hepatitis virus (MHV), a widely used model of beta-CoV, lead to defects in viral replication that can be rescued by compensatory mutations restoring the base-pairing [40]. On the other hand, the lower part of the SL1 needs to be structurally dynamic to establish a transient long-range interaction with the 3′ UTR, enabling the synthesis of subgenomic mRNAs (sgRNAs). Accordingly, studies conducted both in MHV and in human CoV 229E and NL63 (HCoV-229E and HCoV-NL63) showed that mutations in the lower part of the SL1 lead to defective sgRNA synthesis [41]. The SL1 appears to be highly functionally conserved. Indeed, replacement of the MHV SL1 with the SL1 of SARS-CoV has been proven to result in viable, yet more slowly replicating, chimeric viruses [42,43].

The SL2 is probably the most conserved cis-acting structure, as suggested by multiple phylogenetic analyses [32,42,44], hinting at its putative functional relevance. It consists of a highly conserved YUUGY pentaloop, stacked on a 5 bp-long helical stem, folding into a canonical CUYG tetraloop, with the 3′ pyrimidine flipping out of the stack [45]. Mutations disrupting the SL2 stem are not tolerated and lead to impaired sgRNA synthesis in MHV [32]. Accordingly, complete reversion of SL2-disrupted HCoV-229E mutants to the wild-type structure has been observed after just five passages of *in vitro* evolution [41]. Likewise, viability of mutants harboring mutations at G_4 of the pentaloop is severely impaired, and spontaneous mutants of $C_1 > A$ were observed in $G_4 > U$ mutants, suggesting that these two bases are close enough in the loop to base-pair [46]. Mutants in which the SL2 of HCoV-229E has been replaced with the SL2 of either bovine CoV (BCoV) or SARS-CoV were vital and functional comparably to the wild-type virus, supporting the structural and functional conservation of the SL2 [41].

The SL3 (sometimes referred to as SL-II) is conserved only in a small subset of beta and gamma-CoV [44]. When present, this element encompasses the TRS-L. In MHV, although predicted, the existence of SL3 is not supported by *in vitro* Selective 2′ Hydroxyl Acylation analyzed by Primer Extension (SHAPE) analysis [47], with the TRS-L residing in a single-stranded region. In BCoV, HCoV-HKU1, and HCoV-OC43, the SL3 is predicted to form, with the TRS-L residing in the loop. Oppositely, in SARS-CoV and SARS-CoV-2 the TRS-L is predicted to reside in the 3′-half of the stem. This constitutes an exception, as in most CoV the TRS-L is predicted to be single-stranded [48], in agreement with its need to be available for interaction with the TRS-B
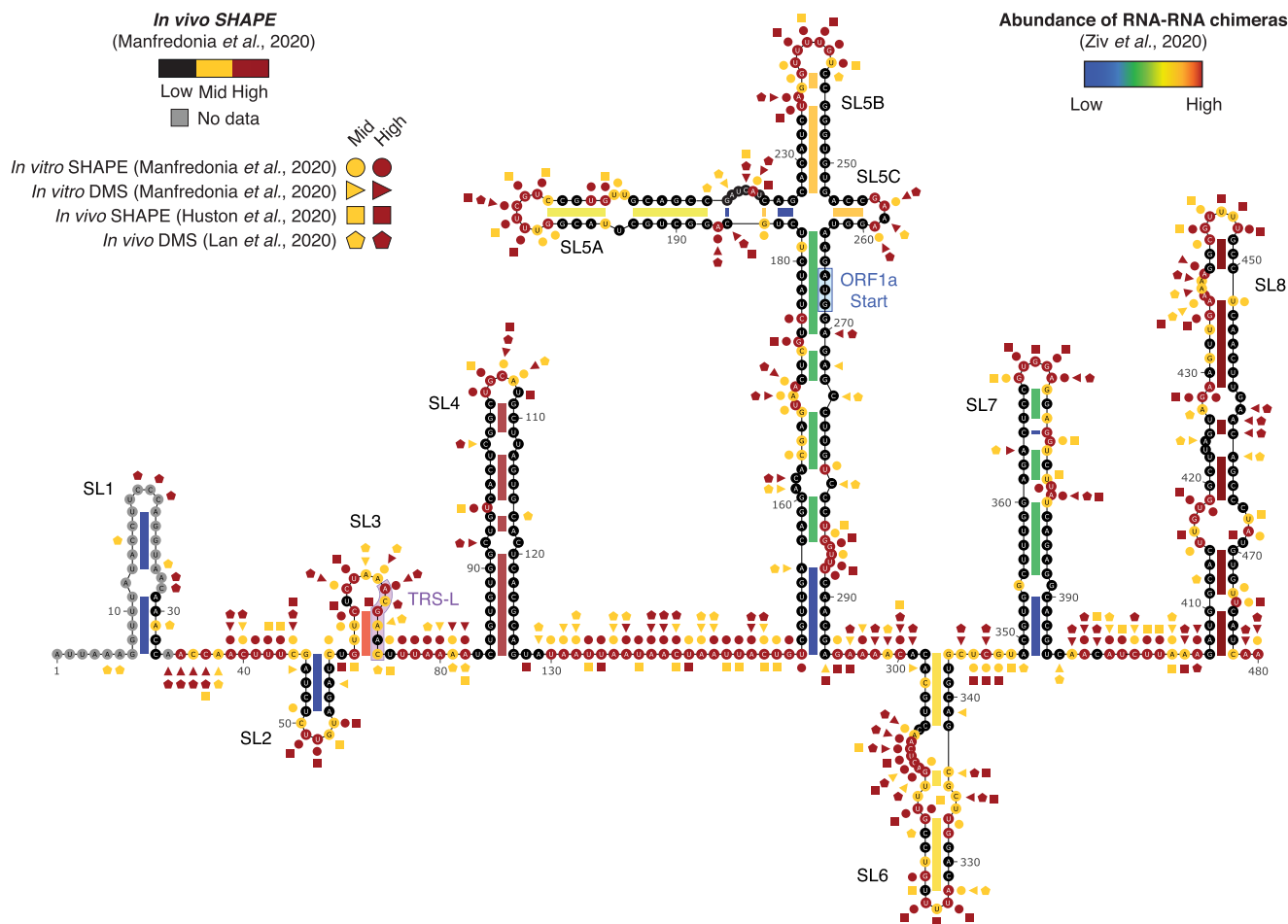
Figure 1. Structure of the SARS-CoV-2 5′ UTR.

Secondary structure of the SARS-CoV-2 5′ UTR, with superimposed *in vivo* SHAPE reactivities from Manfredonia et al. [39]. Highly (red) and moderately (yellow) reactive residues from *in vitro* SHAPE (circles; Manfredonia et al. [39]), *in vitro* DMS (triangles; Manfredonia et al. [39]), *in vivo* SHAPE (squares; Huston et al. [36]) and *in vivo* DMS (pentagons; Lan et al. 2020) experiments are also indicated [36,37,39]. A higher reactivity indicates a higher propensity of bases to be single-stranded (for DMS), or structurally flexible (for SHAPE). Base-paired regions are color-coded according to the number of supporting chimeric reads from Ziv et al. [38]. The number of reads supporting the existence of SL8 was calculated by reanalyzing data from Ziv et al. [38] (GEO dataset: GSE154662).

sequences for proper sgRNA synthesis. Although both *in silico* prediction and *in vitro* SHAPE probing of the SARS-CoV-2 genome support folding of the SL3 [34,39], *in vivo* SHAPE and dimethyl sulfate (DMS) probing revealed intermediate reactivities in the stem of the SL3 [37,39], suggesting that this region might undergo dynamic unfolding to mediate genome cyclization. Direct capture of *in vivo* RNA–RNA interactions further supports this model, revealing the formation of a long-range interaction between the unfolded SL3 segment and the 3′ UTR [38].

The SL4 (sometimes referred to as SL-III) consists of a long bipartite hairpin (SL4a and SL4b, separated by an internal loop) located right downstream of the TRS-L, that appears to be conserved in all CoV genera [44]. This element harbors an upstream ORF (uORF) in over 75% of beta-CoV, encoding for a polypeptide of 3 to 13 amino acids. In MHV, mutations disrupting the coding potential of this uORF while preserving the structure of the SL4 resulted in the increased translation of ORF1ab, and were rapidly spontaneously reverted, although not impairing virus viability [49]. Complete deletion of the uORF start codon, however, resulted in defective viruses, and was spontaneously rescued by the appearance of a new uORF, suggesting that this uORF is involved in regulating the translation of ORF1ab. In MHV, the complete disruption of the SL4 structure, as

well as the separate deletion of SL4a or SL4b, is tolerated, while the complete deletion of the SL4 is lethal [50]. These observations suggest that SL4 might function as a structural spacer, determining the proper orientation of the SL1, SL2, and TRS-L, possibly regulating the synthesis of sgRNAs. Recent *in silico* analysis of the SARS-CoV-2 genome suggested the presence of a second additional stem–loop, right downstream of the SL4 [34], whose existence, however, is supported neither by *in vitro* or *in vivo* structure probing analyses [37,39].

The SL5 partially encompasses the ORF1ab, including the initial portion of the *nsp1* gene, and is conserved in both alpha and beta-CoV [44]. It consists of a four-way junction, comprising three substructural stem–loops, SL5A, SL5B, and SL5C (sometimes referred to as SLIV, SLV, and SLVI). Folding of SL5 has been experimentally confirmed by *in vitro* probing of MHV [47], as well as by both *in vitro* and *in vivo* probing [36,37,39] and by direct *in vivo* RNA–RNA interaction capture in SARS-CoV-2 [38]. While the disruption of SL5A impairs viral replication in MHV, the importance of SL5C is more controversial [51]. Indeed, the deletion of SL5C in MHV is partially tolerated, but its disruption in BCoV viral defective interfering RNAs impairs replication [52]. However, Guan and colleagues pointed out that the mutants designed by Brown and collaborators could potentially affect the formation of the long-range SL5A base-pairing in BCoV, rather than affecting the folding of SL5C. To the best of our knowledge, no study has ever assessed the importance of SL5B. In alpha-CoV, each of the three hairpin loops presents the conserved sequence UUYCGU, while in beta-CoV this sequence is only present in the loops of SL5A and SL5B, and, in most cases (including SARS-CoV and SARS-CoV-2), the SL5C presents a GNRA tetraloop. MERS-CoV constitutes one of the few exceptions, as it harbors a non-conserved heptaloop. As viral encapsidation signals usually involve repeated structural elements, it has been previously proposed that the SL5 might act as a genome packaging signal in CoV [44], although this hypothesis has never been validated by reverse genetics approaches. CoV packaging signals will not be treated here, as they have already been extensively discussed in a recent review [53].

A number of additional SLs have been proposed downstream of the SL5. Although being located downstream of the start codon of ORF1ab, these additional structural elements are usually regarded as part of the 5′ UTR. SL6 and SL7 have been predicted computationally [54], and their existence is supported by *in vitro* probing of MHV [47], as well as by both *in vitro* and *in vivo* probing [33,36,37,39] and by direct *in vivo* RNA–RNA interaction capture in SARS-CoV-2 [38]. These two SLs appear to be less well conserved across beta-CoV [54], and targeted disruption of SL6 by mutagenesis in MHV does not affect viral replication nor viability [47]. Comparative computational analysis of the SARS-CoV-2 genome suggested the presence of three small additional SLs, approximately spanning nucleotides 398 to 450 [34]. However, existence of these SLs was not supported by *in silico* scanning of the SARS-CoV-2 genome in search for structure elements significantly more stable than expected by chance [55], nor by further *in vitro* and *in vivo* structure probing analyses conducted in the context of the full-length SARS-CoV-2 genome [37,39]. Rather, these analyses, as well as *in vivo* RNA–RNA interaction capture experiments [38], suggested the existence of a large SL8 element encompassing this region, spanning nucleotides 407 to 478.

Ribosomal frameshifting element (FSE)

Located at the intersection between ORF1a and ORF1b, the FSE is probably the most well studied structural element in CoV genomes. It regulates the programmed −1 ribosomal frameshifting, enabling the translation of ORF1b, partially overlapping ORF1a. Originally identified in the gamma-CoV infectious bronchitis virus (IBV), and further confirmed in MHV [56], it has been proposed to consist of a slippery UUUAAAC sequence, followed by a single-stranded spacer and a pseudoknotted stimulatory structure [57], composed of two stacked stems (S1 and S2), connected by two large loops (L1 and L2). In SARS-CoV [58] and in SARS-CoV-2 [34], the same overall organization has been suggested, but with a third stem–loop forming within L2 (S3). *In vitro* SHAPE and cryo-EM analyses of the isolated SARS-CoV-2 FSE confirmed this architecture (Figure 2A,B), additionally revealing two alternative tertiary arrangements, one of which shows the 5′ end threading through a ring formed within the pseudoknot, as opposed to an unthreaded conformation [59]. The threaded conformation might be responsible for the ribosomal frameshifting, via a torsional restraint mechanism. In agreement with the observed structure heterogeneity, molecular dynamics simulations of the SARS-CoV-2 FSE revealed multiple alternative packing of helices S1 and S2 [60]. Surprisingly, *in vitro* and *in vivo* DMS and SHAPE probing of SARS-CoV-2 suggest that the three-stem pseudoknot might not represent the predominant conformation in the context of the full genome (Figure 2C). Indeed, SHAPE analysis does not support the existence of helix S1 [39], and DMS-guided analysis of coexisting alternative conformations rather suggests the presence of two mutually-exclusive non-pseudoknotted stem–loop structures [37]. Accordingly, the abundance
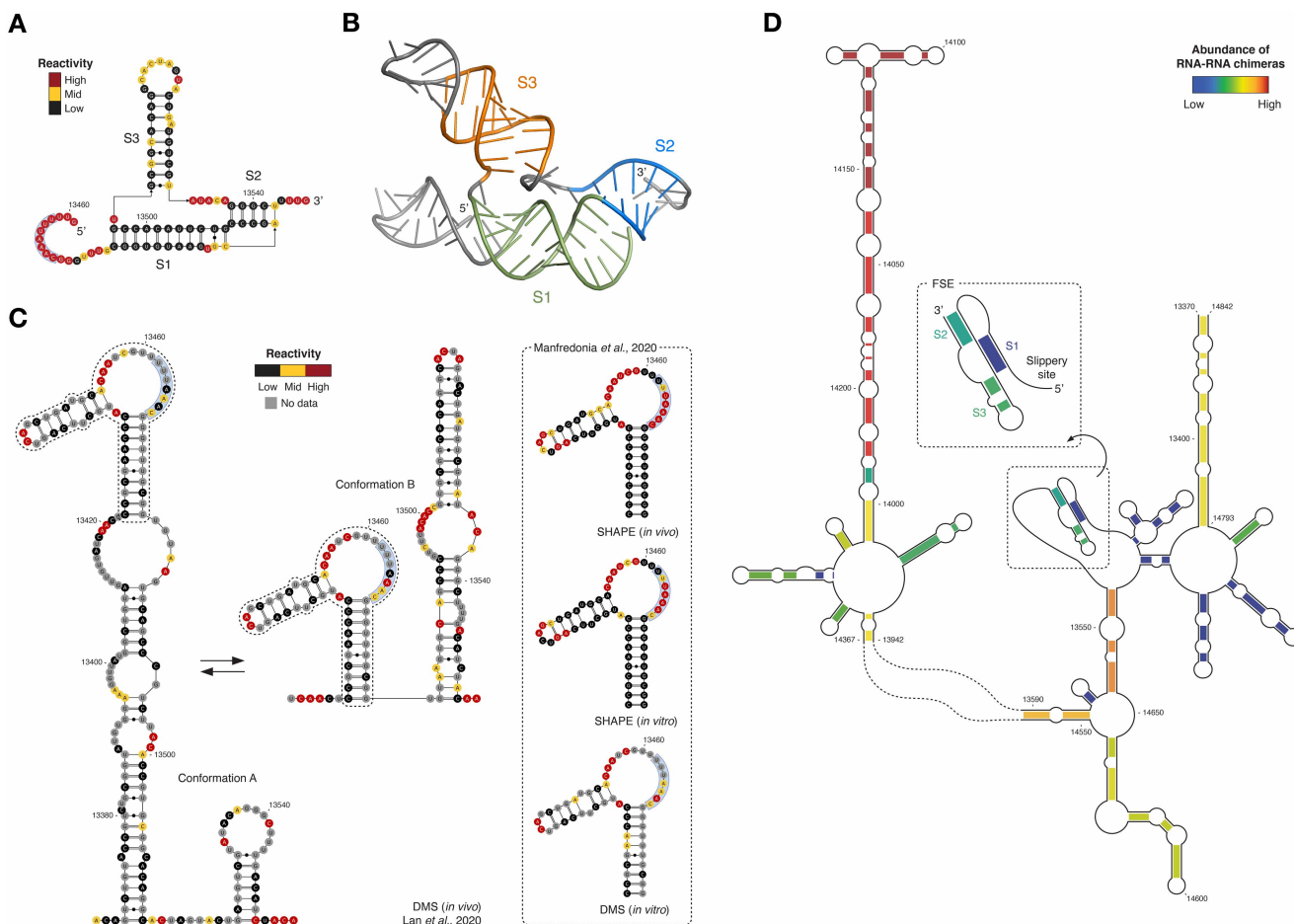
Figure 2. Structure models of the SARS-CoV-2 FSE.
(**A**) Secondary structure of the SARS-CoV-2 FSE three-stem pseudoknotted conformation, with superimposed reactivities from Zhang et al. [59]. (**B**) Cryo-EM-derived structure of the SARS-CoV-2 FSE three-stem pseudoknotted conformation (PDB: 6XRZ). (**C**) On the left, the proposed secondary structure models of two coexisting mutually-exclusive alternative conformations of the SARS-CoV-2 FSE, as derived by *in vivo* DMS analysis, with superimposed reactivities from Lan et al. [37]. The alternative conformation of stem 1 is boxed. On the right, the same alternative conformation of stem 1 as confirmed by independent *in vitro* and *in vivo* DMS and SHAPE analyses, with reactivities superimposed from Manfredonia et al. [39]. The slippery site is boxed in blue. (**D**) Structure of the FSE-arch enclosing the FSE, as identified by direct RNA–RNA interaction mapping in the SARS-CoV-2 genome, with base-pairs colored according to their relative abundance from Ziv et al. [38].

of chimaras from *in vivo* RNA–RNA interaction capture experiments supporting the existence of helix S1 appears to be extremely low [38] (Figure 2D). Furthermore, analysis of RNA–RNA interactions in both SARS-CoV-2 and MERS-CoV revealed that the FSE is embedded within a partially conserved ∼1.5 kb-long higher-order structure, bridging the 3′ end of ORF1a to the 5′ end of ORF1b, dubbed FSE-arch.

3′ UTR structures

Most of the knowledge on the 3′ UTR of coronaviruses comes from studies conducted on beta-CoV. This region appears to be structurally and functionally conserved within the same genus, so that the 3′ UTR of either BCoV or SARS-CoV can functionally replace that of MHV [61,62]. Oppositely, replacement of the MHV 3′ UTR with that of transmissible gastroenteritis virus (TGEV) or of IBV results in non-viable chimaeras, indicating that this functional equivalence is not consistent across different genera.

The most proximal structure element of the 3′ UTR, located immediately downstream of the stop codon of the nucleocapsid (N) gene, is the bulged stem–loop (BSL). This element shows very limited sequence, but high structural conservation across beta-CoV, and it has been proven to be essential for viral replication in MHV

[62,63]. The terminal portion of the BSL, has been proposed to mediate the formation of an alternative, mutually-exclusive pseudoknotted conformation, by direct base-pairing to the loop of a hairpin (P2) located downstream of the BSL [64]. Originally identified in BCoV [65], it was proven to be essential for viral replication. The interconversion between the BSL and the pseudoknot structure has been proposed to act as a molecular switch, regulating viral RNA synthesis, and modulating the transition between different steps of the negative strand synthesis [66]. In alpha-CoV, although the BSL does not seem to be present, the pseudoknot is predicted to exist, and its formation appears to be mutually-exclusive with the presence of a small upstream stem–loop, suggesting that the putative molecular switch proposed in beta-CoV might also exist in alpha-CoV [48]. Thermal denaturation experiments conducted on the MHV 3′ UTR partially supported this model, revealing that the pseudoknot is highly dynamic: it only forms very weakly, and only when the terminal portion of the BSL cannot fold, suggesting that the fully folded BSL might represent the predominant conformation [67]. Accordingly, recent SHAPE-guided structure probing analyses, as well as direct capture of RNA–RNA interactions in the SARS-CoV-2 genome, did not support the presence of this pseudoknot [35,36,38,39]. It is however worth pointing out that the structural analyses performed in all these studies only presented an averaged snapshot of the SARS-CoV-2 RNA structure at a single time-point in a non-synchronous population of infected cells, while the existence of this pseudoknot might be temporally limited to a very specific stage of the viral life cycle.

The P2 stem–loop is part of a large multi-branched structure, known as the hypervariable region (HVR). The HVR shows poor sequence conservation across CoV, except for the octanucleotide GGAAGAGC, that appears to be extremely conserved across all CoV genera [48,68], although its functional relevance still needs to be assessed. It has been originally predicted and biochemically validated in MHV [69]. This region can tolerate extensive mutagenesis, and its complete deletion does not affect viral replication *in vitro* [68]. In contrast, removal of the HVR has a strong impact on viral pathogenicity, as demonstrated by the absence of any clinical sign of infection in mice infected with HVR-deleted viruses. Although the existence of this large structure in the SARS-CoV-2 genome is supported by both *in vitro* structure probing [39] and direct capture of *in vivo* RNA–RNA interactions [38], *in vivo* SHAPE probing reveals extensive unfolding of large part of the main HVR stem [36,39], suggesting that the HVR might adopt multiple alternative conformations *in vivo*.

Even if generally regarded as part of the 3′UTR, the BSL and part of HVR overlap a putative open reading frame (ORF10), whose existence, however, does not appear to be supported by a recent transcriptome analysis of SARS-CoV-2 infected cells [26].

A highly conserved structure element, stemming from the multi-branched loop of the HVR, is the stem–loop II-like motif (s2m). This stem–loop shows extreme sequence and structure similarity to the second stem–loop from the 3′ UTR of astroviruses and equine rhinovirus [70], and its existence in SARS-CoV-2 is supported under both *in vitro* and *in vivo* conditions by multiple studies [33,35–39]. X-ray crystallography of the s2m revealed that the tertiary folding of the apical pentanucleotide GAGUA loop resembles that of a typical GNRA tetraloop, with the U bulging out of the stack [71]. Overall, the s2m has a very particular geometry, forming a sharp 90° kink of the helix axis, resembling the 530 stem–loop of the *Escherichia coli* 16S ribosomal RNA, suggesting that this structure element might be involved in hijacking the host cell protein synthesis machinery.

Other structural elements in the SARS-CoV-2 genome

Besides the well-characterized structures within the 5′ and 3′ UTR, and the FSE, the SARS-CoV-2 genome is predicted to have an exceptionally high propensity to form stable RNA structures, significantly higher compared with that of other RNA viruses, including Hepatitis C virus (HCV), one of the most highly structured viral RNAs to date [72]. Indeed, recent genome-scale structure analyses of the SARS-CoV-2 genome, both *in vitro* and in living infected host cells, have led to the identification of a plethora of novel highly stable RNA structure elements, spread along the whole genome. DMS probing of infected host cells revealed that seven out of the nine TRS-Bs reside within stem–loop structures (Figure 3, top), characterized by different degrees of exposure of the TRS core sequence [37]. Indeed, while in four out of seven stem–loops (M, ORF6, ORF8, and N) the TRS core resides fully or partially within internal loops or single-nucleotide bulges, in three out of seven stem–loops (S, ORF3a, and E) the TRS core resides within base-paired helical regions. It is conceivable that these structures can regulate the accessibility of the TRS-B sequences, to finely tune the synthesis of the different sgRNAs. In line with this hypothesis, *in vivo* SHAPE analysis revealed a correlation between the degree of structuring of the TRS-B, and the relative abundance of the respective sgRNA [35]. Similarly, the
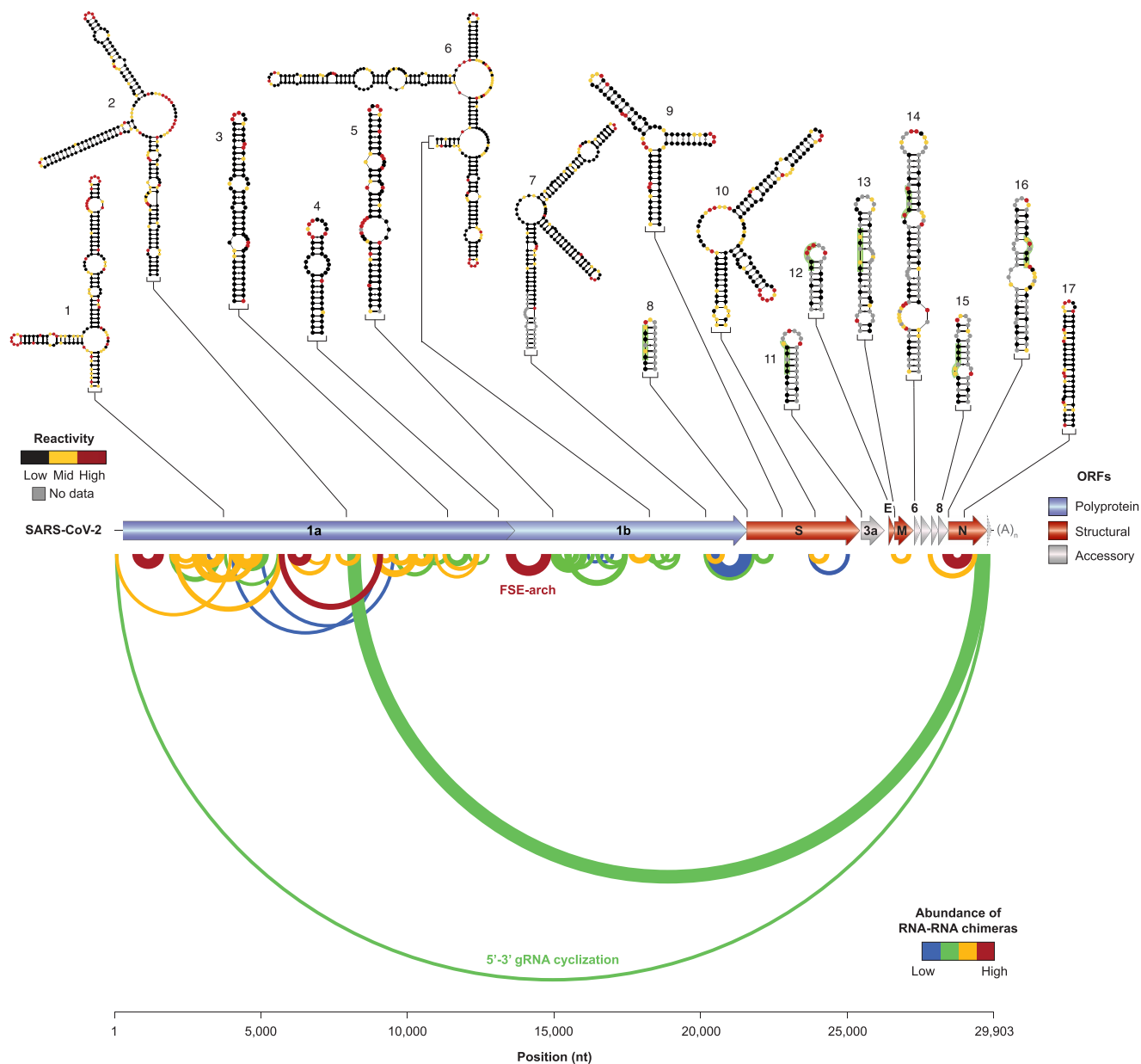
Figure 3. Landscape of SARS-CoV-2 RNA structures as revealed by high-throughput studies.
On the top, the structures of conserved secondary structure elements, supported by significant covariation, as determined by SHAPE analyses of the SARS-CoV-2 genome, with superimposed reactivities from Manfredonia et al. [39] (1, 2, 6, 7, 9, 10, 17) and Huston et al. [36] (3, 4, 5). In addition, the structures of seven stem–loops (8, 11, 12, 13, 14, 15, 16) proposed to enclose the TRS-Bs (boxed in green), with superimposed reactivities from Lan et al. [37], are shown. On the bottom, an arc plot shows the long-range interactions identified by direct RNA–RNA interaction mapping in the SARS-CoV-2 genome, colored according to their relative abundance from Ziv et al. [38]. The FSE-arch is also indicated.

RNA structural features were well correlated with the translation efficiency of sgRNAs, hinting at the importance of RNA structures in regulating the expression levels of viral proteins.

Comparison of the SARS-CoV-2 genome structure modeled by *in vivo* SHAPE analysis, with previously proposed structure models of Dengue (DENV) and HCV viruses, suggested that the SARS-CoV-2 genome is more prone to form locally-stable secondary structures, with fewer long-range base-pairing interactions, as compared with other positive-sense RNA viruses [36]. This preference for locally stable structures over long-range interactions might be involved in safeguarding the stability of SARS-CoV-2 genome and ensuring translation fidelity.

Nonetheless, *in vivo* RNA–RNA interaction capture experiments revealed the existence of several long-range interactions, particularly enriched within ORF1a [38] (Figure 3, bottom). Part of these interactions appears to be mutually-exclusive, suggesting the presence of multiple coexisting alternative conformations. As an example, the 5′ UTR can interact with both the 3′ UTR, possibly to drive genome circularization, and with ORF1a.

Comparative *in vivo* and *in vitro* SHAPE probing analyses of the SARS-CoV-2 genome structure enabled the identification of nearly 87 well-defined structure elements, showing coherent folding under both conditions [39]. Of these, at least 10% showed significant covariation, indicating that they are under strong purifying selection, and appeared to be conserved, to different extents, in alpha-CoV, gamma-CoV, and delta-CoV as well (Figure 3, top). The same study revealed an unexpectedly high degree of correlation between the *in vivo* and *in vitro* conditions, suggesting that the sequence context and thermodynamics alone are major determinants in the folding of the SARS-CoV-2 genome. Accordingly, comparative DMS probing analyses of the FSE *in vivo* and *in vitro* showed that, *in vitro*, folding of the FSE region became progressively more and more similar to that observed *in vivo* when a larger portion of the surrounding sequence context was included in the *in vitro* transcript, with the DMS patterns being essentially indistinguishable when *in vitro* refolding the full SARS-CoV-2 genome [37].

Several studies have further started investigating the presence of RNA G-quadruplexes (G4s) in the SARS-CoV-2 genome [73–76]. Of the many predicted G4s, one G4 located within ORF1a (position 13385) was shown to bind and to be stabilized by BRACO-19 and TMPyP4, two known G4 binders, suggesting that these compounds might represent good scaffolds for the development of RNA-targeted small-molecule drugs [74]. Nevertheless, currently available studies only analyzed the candidate G4s under *in vitro* conditions. Further analyses will be needed to determine whether G4s can form in the SARS-CoV-2 genome, in the context of the infected cell.

Additional insights from transcriptome-wide studies

Apart from intramolecular base-pairings, the SARS-CoV-2 genome appears to be involved also in a number of intermolecular RNA–RNA interactions with the host cell's transcriptome [38]. Particularly, direct capture of *in vivo* RNA–RNA interactions showed that the SARS-CoV-2 gRNA establishes several interactions with the small nuclear RNAs (snRNAs) U1, U2 and U4 at the level of ORF1ab, while the remainder of the gRNA is devoid of such interactions. Oppositely, the sgRNAs are enriched for interactions with the U1 and U2 snRNAs within the N gene and the 3′ UTR. Interactions with the U2 snRNA within ORF1a appear to be highly conserved, as confirmed by RNA–RNA interaction mapping in MERS-CoV-infected cells.

Additionally, gRNA pulldown experiments coupled to mass-spectrometry analyses showed that, during the SARS-CoV-2 life cycle, the gRNA establishes numerous RNA–protein interactions, both with viral and host cell proteins [30,31,77]. Among the top-enriched interactors is the SARS-CoV-2 N protein. Multiple computational and experimental analyses provided compelling evidences that the N protein has a high propensity to undergo liquid–liquid phase separation (LLPS), that is stimulated by interaction with the viral gRNA [78–81]. This interaction preferentially occurs with single-stranded RNA regions flanked by stably structured elements [82], further suggesting that LLPS might be exploited by SARS-CoV-2 for efficient genome packaging. Of the several interacting host cell proteins identified, the vast majority appears to have a host-protective effect, as their knockout resulted into virus-induced cell death [31], and to be part of the RNA–protein interactome of other positive-sense RNA viruses such as Zika virus (ZIKV) and DENV [30,31]. Of these, CNBP, a zinc-finger protein able to activate the expression of several innate immune response genes, was confirmed to directly bind the SARS-CoV-2 genome by enhanced cross-linking and immunoprecipitation (eCLIP) [30], suggesting that it might be involved in sensing foreign RNA. SHFL (also known as RyDEN), previously reported to inhibit the programmed -1 ribosomal frameshifting in human immunodeficiency virus (HIV) infections [83], was further shown to be able to inhibit ribosomal frameshifting by the SARS-CoV-2 FSE in a dual-color fluorescence reporter system. Interestingly, while the ZIKV and DENV genomes were found to be robustly associated with N6-methyladenosine (m^6A) readers of the YTHDF family, and depleted of interactions with m^6A erasers of the ALKBH family, the SARS-CoV-2 genome showed the opposite trend, suggesting that SARS-CoV-2 might escape m^6A methylation to increase the gRNA stability [31]. Oppositely, a specific interaction with the mitochondrial 2′-O-methyltransferase MRM2 was observed, in agreement with the physical localization of the SARS-CoV-2 gRNA to mitochondria, as suggested by a recent machine learning analysis revealing the presence of several mitochondrial-localization signals in the 5′ UTR of the SARS-CoV-2 genome [84]. Direct RNA sequencing of the SARS-CoV-2 gRNA and sgRNAs further supports the notion that these RNAs might be

post-transcriptionally modified [26,85]. Although the exact nature of these modifications is still unclear, a recent study suggests that at least 42 5-methylcytosine (m^{5}C) sites might exist in SARS-CoV-2 RNAs [86].

Concluding remarks

Recent advances in the RNA field have proven that RNA structure represents an ideal, yet largely under-exploited, drug target [87,88]. In this perspective, the presence of a plethora of structured RNA elements in coronavirus genomes provides a unique opportunity for the development of new effective therapeutic strategies. Over the past few months, unprecedented efforts have been put in trying to dissect the complexity of the SARS-CoV-2 genome, revealing a multitude of previously unannotated RNA structure elements. Although their role has yet to be assessed, the conservation of a subset of these structures hints at their functional relevance. Importantly, preliminary analyses suggest that some of these elements might present druggable pockets [39], and early attempts to develop RNA-targeted therapeutic strategies are already underway [89].

Perspectives

- Given the fast evolution of RNA viruses, leading to changes in the sequence and structure of viral proteins, the existence of conserved RNA structures in viral genomes provides a unique opportunity to develop more potent and durable antiviral therapeutic strategies

- Multiple groups have queried the structure of the SARS-CoV-2 genome under different conditions (*in vitro* or *in vivo*), using orthogonal probing approaches (DMS, NAI, NAI-N3), and reported diverse sets of RNA structures, providing a wide repertoire of potential therapeutic targets

- While most studies so far only provided a static snapshot of RNA structures in CoV genomes (at a single time-point of infection), future efforts should be aimed at characterizing the dynamics of the CoV RNA structurome to possibly identify crucial transient RNA folds and RNA structure switches, whose targeting might provide the means to effectively inhibit viral replication

Abbreviations

CoV, coronavirus; DMS, dimethyl sulfate; gRNA, genomic RNA; MERS, Middle-East respiratory syndrome; SARS, severe acute respiratory syndrome; sgRNA, subgenomic mRNA; SHAPE, Selective 2′ Hydroxyl Acylation analyzed by Primer Extension.

References

1 Carrasco-Hernandez, R., Jácome, R., Vidal, Y.L. and de León, S.P. (2017) Are RNA viruses candidate agents for the next global pandemic? a review. *Ilar. J.* **58**, 343–358 https://doi.org/10.1093/ilar/ilx026

2 Woolhouse, M.E.J., Adair, K. and Brierley, L. (2014) RNA viruses: a case study of the biology of emerging infectious diseases. *Microbiol. Spectr.* **1**, 83–97 https://doi.org/10.1128/microbiolspec.OH-0001-2012

3 Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. and VandePol, S. (1982) Rapid evolution of RNA genomes. *Science* **215**, 1577–1585 https://doi.org/10.1126/science.7041255

4 Choi, K.H. (2011) Viral Polymerases. *Adv. Exp. Med. Biol.* **726**, 267–304 https://doi.org/10.1007/978-1-4614-0980-9_12

5 Boerneke, M.A., Ehrhardt, J.E. and Weeks, K.M. (2019) Physical and functional analysis of viral RNA genomes by SHAPE. *Ann. Rev. Virol.* **6**, 93–117 https://doi.org/10.1146/annurev-virology-092917-043315

6 Ouellet, D.L., Plante, I., Landry, P., Barat, C., Janelle, M.-È., Flamand, L. et al. (2008) Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res.* **36**, 2353–2365 https://doi.org/10.1093/nar/gkn076

7 Bannwarth, S. and Gatignol, A. (2005) HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr. Hiv Res.* **3**, 61–71 https://doi.org/10.2174/1570162052772924

8 Roy, S., Delling, U., Chen, C.H., Rosen, C.A. and Sonenberg, N. (1990) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Gene Dev.* **4**, 1365–1373 https://doi.org/10.1101/gad.4.8.1365

9 Huthoff, H. and Berkhout, B. (2001) Two alternating structures of the HIV-1 leader RNA. *RNA* **7**, 143–157 https://doi.org/10.1017/S1355838201001881

10 Rausch, J.W., Sztuba-Solinska, J. and Grice, S.F.J.L. (2018) Probing the structures of viral RNA regulatory elements with SHAPE and related methodologies. *Front. Microbiol.* **8**, 2634 https://doi.org/10.3389/fmicb.2017.02634

11 Incarnato, D. and Oliviero, S. (2017) The RNA epistructurome: uncovering RNA function by studying structure and post-transcriptional modifications. *Trends Biotechnol.* **35**, 318–333 https://doi.org/10.1016/j.tibtech.2016.11.002

12 Tomezsko, P.J., Corbin, V.D.A., Gupta, P., Swaminathan, H., Glasgow, M., Persad, S. et al. (2020) Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **582**, 438–442 https://doi.org/10.1038/s41586-020-2253-5

13 Cavanagh, D. (2005) Coronaviridae: a review of coronaviruses and toroviruses. In *Coronaviruses with Special Emphasis on First Insights Concerning SARS. Birkhäuser Advances in Infectious Diseases BAID* (Schmidt, A., Weber, O., Wolff, M.H., eds), pp. 1–54. Birkhäuser Basel

14 Cui, J., Li, F. and Shi, Z.-L. (2018) Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 https://doi.org/10.1038/s41579-018-0118-9

15 de Wit, E., van Doremalen, N., Falzarano, D. and Munster, V.J. (2016) SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 https://doi.org/10.1038/nrmicro.2016.81

16 Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W. et al. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 https://doi.org/10.1038/s41586-020-2012-7

17 Lauber, C., Goeman, J.J., del C Parquet, M., Nga, P.T., Snijder, E.J., Morita, K. et al. (2013) The footprint of genome architecture in the largest genome expansion in RNA viruses. *PLoS Pathog.* **9**, e1003500 https://doi.org/10.1371/journal.ppat.1003500

18 Minskaia, E., Hertzig, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B. et al. (2006) Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl Acad. Sci. U.S.A.* **103**, 5108–5113 https://doi.org/10.1073/pnas.0508200103

19 Ferron, F., Subissi, L., Morais, A.T.S.D., Le, N.T.T., Sevajol, M., Gluais, L. et al. (2018) Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl Acad. Sci. U.S.A.* **115**, E162–E171 https://doi.org/10.1073/pnas.1718806115

20 Sawicki, S.G., Sawicki, D.L. and Siddell, S.G. (2007) A contemporary view of coronavirus transcription. *J. Virol.* **81**, 20–29 https://doi.org/10.1128/JVI.01358-06

21 Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L. et al. (2020) Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **368**, 779–782 https://doi.org/10.1126/science.abb7498

22 Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T. and Veesler, D. (2020) Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6 https://doi.org/10.1016/j.cell.2020.02.058

23 Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y. and Zhou, Q. (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448 https://doi.org/10.1126/science.abb2762

24 Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S. et al. (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 https://doi.org/10.1038/s41586-020-2180-5

25 Krafcikova, P., Silhan, J., Nencka, R. and Boura, E. (2020) Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat. Commun.* **11**, 3717 https://doi.org/10.1038/s41467-020-17495-9

26 Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J.W., Kim, V.N. and Chang, H. (2020) The Architecture of SARS-CoV-2 transcriptome. *Cell* **181**, 914–921.e10 https://doi.org/10.1016/j.cell.2020.04.011

27 Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M. et al. (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 https://doi.org/10.1038/s41586-020-2286-9

28 Sadegh, S., Matschinske, J., Blumenthal, D.B., Galindez, G., Kacprowski, T., List, M. et al. (2020) Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. *Nat. Commun.* **11**, 3518 https://doi.org/10.1038/s41467-020-17189-2

29 Perrin-Cocon, L., Diaz, O., Jacquemin, C., Barthel, V., Ogire, E., Ramière, C. et al. (2020) The current landscape of coronavirus-host protein–protein interactions. *J. Transl. Med.* **18**, 319 https://doi.org/10.1186/s12967-020-02480-z

30 Schmidt, N., Lareau, C.A., Keshishian, H., Melanson, R., Zimmer, M., Kirschner, L. et al. (2020) A direct RNA-protein interaction atlas of the SARS-CoV-2 RNA in infected human cells. *Biorxiv* https://doi.org/10.1101/2020.07.15.204404

31 Flynn, R.A., Belk, J.A., Qi, Y., Yasumoto, Y., Schmitz, C.O., Mumbach, M.R. et al. (2020) Systematic discovery and functional interrogation of SARS-CoV-2 viral RNA-host protein interactions during infection. *Biorxiv* https://doi.org/10.1101/2020.10.06.327445

32 Liu, P., Li, L., Millership, J.J., Kang, H., Leibowitz, J.L. and Giedroc, D.P. (2007) A U-turn motif-containing stem-loop in the coronavirus 5′ untranslated region plays a functional role in replication. *RNA* **13**, 763–780 https://doi.org/10.1261/rna.261807

33 Sanders, W., Fritch, E.J., Madden, E.A., Graham, R.L., Vincent, H.A., Heise, M.T. et al. (2020) Comparative analysis of coronavirus genomic RNA structure reveals conservation in SARS-like coronaviruses. *Biorxiv* https://doi.org/10.1101/2020.06.15.153197

34 Rangan, R., Zheludev, I.N. and Das, R. (2020) RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *RNA* **26**, 937–959 https://doi.org/10.1261/rna.076141.120

35 Sun, L., Li, P., Ju, X., Rao, J., Huang, W., Zhang, S. et al. (2020) In vivo structural characterization of the whole SARS-CoV-2 RNA genome identifies host cell target proteins vulnerable to re-purposed drugs. *Biorxiv* https://doi.org/10.1101/2020.07.07.192732

36 Huston, N.C., Wan, H., de CA Tavares, R., Wilen, C. and Pyle, A.M. (2020) Comprehensive in-vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Biorxiv* https://doi.org/10.1101/2020.07.10.197079

37 Lan, T.C.T., Allan, M.F., Malsick, L.E., Khandwala, S., Nyeo, S.S.Y., Bathe, M. et al. (2020) Structure of the full SARS-CoV-2 RNA genome in infected cells. *Biorxiv* https://doi.org/10.1101/2020.06.29.178343

38 Ziv, O., Price, J., Shalamova, L., Kamenova, T., Goodfellow, I., Weber, F. et al. (2020) The short- and long-range RNA-RNA Interactome of SARS-CoV-2. *Mol. Cell* **80**, 1067–1077.e5 https://doi.org/10.1016/j.molcel.2020.11.004

39 Manfredonia, I., Nithin, C., Ponce-Salvatierra, A., Ghosh, P., Wirecki, T.K., Marinus, T. et al. (2020) Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.* **48**, 12436–12452 https://doi.org/10.1093/nar/gkaa1053

40 Li, L., Kang, H., Liu, P., Makkinje, N., Williamson, S.T., Leibowitz, J.L. et al. (2008) Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. *J. Mol. Biol.* **377**, 790–803 https://doi.org/10.1016/j.jmb.2008.01.068

41 Madhugiri, R., Karl, N., Petersen, D., Lamkiewicz, K., Fricke, M., Wend, U. et al. (2018) Structural and functional conservation of cis-acting RNA elements in coronavirus 5′-terminal genome regions. *Virology* **517**, 44–55 https://doi.org/10.1016/j.virol.2017.11.025

42 Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P. and Leibowitz, J.L. (2006) Putative cis-acting stem-loops in the 5′ untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. *J. Virol.* **80**, 10600–10614 https://doi.org/10.1128/JVI.00455-06

43 Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P. and Leibowitz, J.L. (2006) Stem-loop 1 in the 5′ UTR of the SARS coronavirus can substitute for its counterpart in mouse hepatitis virus. *Adv. Exp. Med Biol.* **581**, 105–108 https://doi.org/10.1007/978-0-387-33012-9_18

44 Chen, S.-C. and Olsthoorn, R.C.L. (2010) Group-specific structural features of the 5′-proximal sequences of coronavirus genomic RNAs. *Virology* **401**, 29–41 https://doi.org/10.1016/j.virol.2010.02.007

45 Lee, C.W., Li, L. and Giedroc, D.P. (2011) The solution structure of coronaviral stem-loop 2 (SL2) reveals a canonical CUYG tetraloop fold. *FEBS Lett.* **585**, 1049–1053 https://doi.org/10.1016/j.febslet.2011.03.002

46 Liu, P., Li, L., Keane, S.C., Yang, D., Leibowitz, J.L. and Giedroc, D.P. (2009) Mouse hepatitis virus stem-loop 2 adopts a uYNMG(U)a-like tetraloop structure that is highly functionally tolerant of base substitutions. *J. Virol.* **83**, 12084–12093 https://doi.org/10.1128/JVI.00915-09

47 Yang, D., Liu, P., Wudeck, E.V., Giedroc, D.P. and Leibowitz, J.L. (2014) SHAPE analysis of the RNA secondary structure of the Mouse Hepatitis Virus 5′ untranslated region and N-terminal nsp1 coding sequences. *Virology* **475**, 15–27 https://doi.org/10.1016/j.virol.2014.11.001

48 Madhugiri, R., Fricke, M., Marz, M. and Ziebuhr, J. (2014) RNA structure analysis of alphacoronavirus terminal genome regions. *Virus Res.* **194**, 76–89 https://doi.org/10.1016/j.virusres.2014.10.001

49 Wu, H.-Y., Guan, B.-J., Su, Y.-P., Fan, Y.-H. and Brian, D.A. (2014) Reselection of a Genomic Upstream Open Reading Frame in Mouse Hepatitis Coronavirus 5′-untranslated-Region Mutants. *J. Virol.* **88**, 846–858 https://doi.org/10.1128/JVI.02831-13

50 Yang, D., Liu, P., Giedroc, D.P. and Leibowitz, J. (2011) Mouse hepatitis virus stem-loop 4 functions as a spacer element required to drive subgenomic RNA synthesis. *J. Virol.* **85**, 9199–9209 https://doi.org/10.1128/JVI.05092-11

51 Guan, B.-J., Su, Y.-P., Wu, H.-Y. and Brian, D.A. (2012) Genetic evidence of a long-range RNA-RNA interaction between the genomic 5′ untranslated region and the nonstructural protein 1 coding region in murine and bovine coronaviruses. *J. Virol.* **86**, 4631–4643 https://doi.org/10.1128/JVI.06265-11

52 Brown, C.G., Nixon, K.S., Senanayake, S.D. and Brian, D.A. (2007) An RNA stem-loop within the bovine coronavirus nsp1 coding region is a cis-acting element in defective interfering RNA replication. *J. Virol.* **81**, 7716–7724 https://doi.org/10.1128/JVI.00549-07

53 Masters, P.S. (2019) Coronavirus genomic RNA packaging. *Virology* **537**, 198–207 https://doi.org/10.1016/j.virol.2019.08.031

54 Madhugiri, R., Fricke, M., Marz, M. and Ziebuhr, J. (2016) Coronavirus cis-acting RNA elements. *Adv Virus Res.* **96**, 127–163 https://doi.org/10.1016/bs.aivir.2016.08.007

55 Andrews, R.J., Peterson, J.M., Haniff, H.F., Chen, J., Williams, C., Greffe, M. et al. (2020) An in silico map of the SARS-CoV-2 RNA Structurome. *Biorxiv* https://doi.org/10.1101/2020.04.17.045161

56 Bredenbeek, P.J., Pachuk, C.J., Noten, A.F.H., Charité, J., Luytjes, W., Weiss, S.R. et al. (1990) The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res.* **18**, 1825–1832 https://doi.org/10.1093/nar/18.7.1825

57 Brierley, I., Digard, P. and Inglis, S.C. (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* **57**, 537–547 https://doi.org/10.1016/0092-8674(89)90124-4

58 Plant, E.P., Pérez-Alvarado, G.C., Jacobs, J.L., Mukhopadhyay, B., Hennig, M. and Dinman, J.D. (2005) A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.* **3**, e172 https://doi.org/10.1371/journal.pbio.0030172

59 Zhang, K., Zheludev, I.N., Hagey, R.J., Wu, M.T.-P., Haslecker, R., Hou, Y.J. et al. (2020) Cryo-electron microscopy and exploratory antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Biorxiv* https://doi.org/10.1101/2020.07.18.209270

60 Omar, S.I., Zhao, M., Sekar, R.V., Moghadam, S.A., Tuszynski, J.A. and Woodside, M.T. (2020) Modeling the structure of the frameshift stimulatory pseudoknot in SARS-CoV-2 reveals multiple possible conformers. *Biorxiv* https://doi.org/10.1101/2020.06.08.141150

351

61 Goebel, S.J., Taylor, J. and Masters, P.S. (2004) The 3′ cis-acting genomic replication element of the severe acute respiratory syndrome coronavirus can function in the murine coronavirus genome. *J. Virol.* **78**, 7846–7851 https://doi.org/10.1128/JVI.78.14.7846-7851.2004

62 Hsue, B. and Masters, P.S. (1997) A bulged stem-loop structure in the 3′ untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J. Virol.* **71**, 7567–7578 https://doi.org/10.1128/JVI.71.10.7567-7578.1997

63 Hsue, B., Hartshorne, T. and Masters, P.S. (2000) Characterization of an essential RNA secondary structure in the 3′ untranslated region of the murine coronavirus genome. *J. Virol.* **74**, 6911–6921 https://doi.org/10.1128/JVI.74.15.6911-6921.2000

64 Yang, D. and Leibowitz, J.L. (2015) The structure and functions of coronavirus genomic 3′ and 5′ ends. *Virus Res.* **206**, 120–133 https://doi.org/10.1016/j.virusres.2015.02.025

65 Williams, G.D., Chang, R.Y. and Brian, D.A. (1999) A phylogenetically conserved hairpin-type 3′ untranslated region pseudoknot functions in coronavirus RNA replication. *J. Virol.* **73**, 8349–8355 https://doi.org/10.1128/JVI.73.10.8349-8355.1999

66 Goebel, S.J., Hsue, B., Dombrowski, T.F. and Masters, P.S. (2004) Characterization of the RNA components of a putative molecular switch in the 3′ untranslated region of the murine coronavirus genome. *J. Virol.* **78**, 669–682 https://doi.org/10.1128/JVI.78.2.669-682.2004

67 Stammler, S.N., Cao, S., Chen, S.-J. and Giedroc, D.P. (2011) A conserved RNA pseudoknot in a putative molecular switch domain of the 3′-untranslated region of coronaviruses is only marginally stable. *RNA* **17**, 1747–1759 https://doi.org/10.1261/rna.2816711

68 Goebel, S.J., Miller, T.B., Bennett, C.J., Bernard, K.A. and Masters, P.S. (2007) A hypervariable region within the 3′ cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. *J Virol.* **81**, 1274–1287 https://doi.org/10.1128/JVI.00803-06

69 Liu, Q., Johnson, R.F. and Leibowitz, J.L. (2001) Secondary structural elements within the 3′ untranslated region of mouse hepatitis virus strain JHM genomic RNA. *J. Virol.* **75**, 12105–12113 https://doi.org/10.1128/JVI.75.24.12105-12113.2001

70 Jonassen, C.M., Jonassen, T.O. and Grinde, B. (1998) A common RNA motif in the 3′ end of the genomes of astroviruses, avian infectious bronchitis virus and an equine rhinovirus. *J. Gen. Virol.* **79**, 715–718 https://doi.org/10.1099/0022-1317-79-4-715

71 Robertson, M.P., Igel, H., Baertsch, R., Haussler, D., Ares, M. and Scott, W.G. (2005) The structure of a rigorously conserved RNA element within the SARS virus genome. *PLoS Biol.* **3**, e5 https://doi.org/10.1371/journal.pbio.0030005

72 de CA Tavares, R., Mahadeshwar, G., Wan, H., QHuston, N.C. and Pyle, A.M. (2020) The global and local distribution of RNA structure throughout the SARS-CoV-2 genome. *J Virol* https://doi.org/10.1128/JVI.02190-20

73 Zhang, R., Xiao, K., Gu, Y., Liu, H. and Sun, X. (2020) Whole genome identification of potential G-quadruplexes and analysis of the G-quadruplex binding domain for SARS-CoV-2. *Front. Genet.* **11**, 587829 https://doi.org/10.3389/fgene.2020.587829

74 Cui, H. and Zhang, L. (2020) G-quadruplexes are present in human coronaviruses including SARS-CoV-2. *Front. Microbiol.* **11**, 567317 https://doi.org/10.3389/fmicb.2020.567317

75 Ji, D., Juhas, M., Tsang, C.M., Kwok, C.K., Li, Y. and Zhang, Y. (2020) Discovery of G-quadruplex-forming sequences in SARS-CoV-2. *Brief Bioinform.*, bbaa114 https://doi.org/10.1093/bib/bbaa114

76 Zhao, C., Qin, G., Niu, J., Wang, Z., Wang, C., Ren, J. et al. (2020) Targeting RNA G-Quadruplex in SARS-CoV-2: a Promising Therapeutic Target for COVID-19? *Angew. Chem. Int. Ed. Engl.* https://doi.org/10.1002/anie.202011419

77 Kamel, W., Noerenberg, M., Cerikan, B., Chen, H., Järvelin, A.I., Kammoun, M. et al. (2020) Global analysis of protein-RNA interactions in SARS-CoV-2 infected cells reveals key regulators of infection. *Biorxiv* https://doi.org/10.1101/2020.11.25.398008

78 Cascarina, S.M. and Ross, E.D. (2020) A proposed role for the SARS-CoV-2 nucleocapsid protein in the formation and regulation of biomolecular condensates. *FASEB J.* **34**, 9832–9842 https://doi.org/10.1096/fj.202001351

79 Vandelli, A., Monti, M., Milanetti, E., Armaos, A., Rupert, J., Zacco, E. et al. (2020) Structural analysis of SARS-CoV-2 genome and predictions of the human interactome. *Nucleic Acids Res.* **48**, 11270–11283 https://doi.org/10.1093/nar/gkaa864

80 Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward, M.D. et al. (2020) The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Biorxiv* https://doi.org/10.1101/2020.06.17.158121

81 Perdikari, T.M., Murthy, A.C., Ryan, V.H., Watters, S., Naik, M.T. and Fawzi, N.L. (2020) SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *EMBO J.* **39**, e106478 https://doi.org/10.15252/embj.2020106478

82 Iserman, C., Roden, C.A., Boerneke, M.A., Sealfon, R.S.G., McLaughlin, G.A., Jungreis, I. et al. (2020) Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell* **80**, 1078–1091.e6 https://doi.org/10.1016/j.molcel.2020.11.041

83 Wang, X., Xuan, Y., Han, Y., Ding, X., Ye, K., Yang, F. et al. (2019) Regulation of HIV-1 Gag-Pol expression by shiftless, an inhibitor of programmed −1 ribosomal frameshifting. *Cell* **176**, 625–635.e14 https://doi.org/10.1016/j.cell.2018.12.030

84 Wu, K.E., Fazal, F.M., Parker, K.R., Zou, J. and Chang, H.Y. (2020) RNA-GPS predicts SARS-CoV-2 RNA residency to host mitochondria and nucleolus. *Cell Syst.* **11**, 102–108.e3 https://doi.org/10.1016/j.cels.2020.06.008

85 Miladi, M., Fuchs, J., Maier, W., Weigang, S., Pedrosa, N.D., Weiss, L. et al. (2020) The landscape of SARS-CoV-2 RNA modifications. *Biorxiv* https://doi.org/10.1101/2020.07.18.204362

86 Taiaroa, G., Rawlinson, D., Featherstone, L., Pitt, M., Caly, L., Druce, J. et al. (2020) Direct RNA sequencing and early evolution of SARS-CoV-2. *Biorxiv* https://doi.org/10.1101/2020.03.05.976167

87 Warner, K.D., Hajdin, C.E. and Weeks, K.M. (2018) Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **17**, 547–558 https://doi.org/10.1038/nrd.2018.93

88 Connelly, C.M., Moon, M.H. and Schneekloth, J.S. (2016) The emerging role of RNA as a therapeutic target for small molecules. *Cell Chem. Biol.* **23**, 1077–1090 https://doi.org/10.1016/j.chembiol.2016.05.021

89 Haniff, H.S., Tong, Y., Liu, X., Chen, J.L., Suresh, B.M., Andrews, R.J. et al. (2020) Targeting the SARS-CoV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RIBOTAC) degraders. *ACS Cent. Sci.* **6**, 1713–1721 https://doi.org/10.1021/acscentsci.0c00984