



Tertiary folds of the SL5 RNA from the 5' proximal region of SARS-CoV-2 and related coronaviruses

Rachael C. Kretsch^a, Lily Xu^b, Ivan N. Zheludev^c, Xueting Zhou^d, Rui Huang^c, Grace Nye^c, Shanshan Li^e, Kaiming Zhang^e, Wah Chiu^{a,d,f,1}, and Rhiju Das^{a,c,g,1}

Contributed by Wah Chiu; received November 26, 2023; accepted January 5, 2024; reviewed by Hong-Wei Wang and Sarah A. Woodson

Coronavirus genomes sequester their start codons within stem-loop 5 (SL5), a structured, 5' genomic RNA element. In most alpha- and betacoronaviruses, the secondary structure of SL5 is predicted to contain a four-way junction of helical stems, some of which are capped with UUYUGU hexaloops. Here, using cryogenic electron microscopy (cryo-EM) and computational modeling with biochemically determined secondary structures, we present three-dimensional structures of SL5 from six coronaviruses. The SL5 domain of betacoronavirus severe-acute-respiratory-syndrome-related coronavirus 2 (SARS-CoV-2), resolved at 4.7 Å resolution, exhibits a T-shaped structure, with its UUYUGU hexaloops at opposing ends of a coaxial stack, the T's "arms." Further analysis of SL5 domains from SARS-CoV-1 and MERS (7.1 and 6.4 to 6.9 Å resolution, respectively) indicate that the junction geometry and inter-hexal loop distances are conserved features across these human-infecting betacoronaviruses. The MERS SL5 domain displays an additional tertiary interaction, which is also observed in the non-human-infecting betacoronavirus BtCoV-HKU5 (5.9 to 8.0 Å resolution). SL5s from human-infecting alphacoronaviruses, HCoV-229E and HCoV-NL63 (6.5 and 8.4 to 9.0 Å resolution, respectively), exhibit the same coaxial stacks, including the UUYUGU-capped arms, but with a phylogenetically distinct crossing angle, an X-shape. As such, all SL5 domains studied herein fold into stable tertiary structures with cross-genus similarities and notable differences, with implications for potential protein-binding modes and therapeutic targets.

cryo-EM | comparative structural biology | coronaviruses | modeling | viral RNA structure

In the *Coronaviridae* family, seven species, SARS-CoV-2, SARS-CoV-1, MERS, HCoV-HKU1, HCoV-OC43, HCoV-229E, and HCoV-NL63, are known to infect humans, and all are derived from the alpha- and betacoronavirus genera (1). While effective vaccines are available against SARS-CoV-2, the likelihood of future coronavirus pandemics motivates efforts to understand the conserved features of coronavirus machinery. The SARS-CoV-2 proteome and its interactions with host proteins have been well studied structurally, enabling the design of therapeutics targeting host and viral proteins (2–4), but the RNA genome has been relatively understudied. On one hand, the monumental scientific response to COVID-19 has resulted in extensive mapping of SARS-CoV-2 RNA secondary structure (5–10). On the other hand, to date, only a few small fragments of the RNA genome—stem-loop 1 in complex with non-structural protein 1, stem-loop 2, and stem-loop 4 from the 5' UTR; the frameshift stimulation element; and the stem-loop 2 motif from the 3' UTR—have been characterized in three dimensions (11–18). While current designs of RNA-targeting therapeutics are often based on the RNA's two-dimensional (2D) base-pairing pattern, known as RNA secondary structure (12, 19–21), three-dimensional (3D) structures are necessary for structure-guided design of many classes of therapeutics (5, 22, 23).

The 5' proximal region of the SARS-CoV-2 genome is a highly structured and functionally important genomic locus (24, 25). This region is divided into secondary structure domains termed "stem-loops," with multiple stem-loops predicted to be present in the 5' proximal region across the coronavirus family (25–27). Stem-loop 5 (SL5, residues 150 to 294) contains the start codon of open reading frame 1a/b (ORF1a/b, residues 266 to 268). Additionally, phylogenetic covariance analysis and chemical probing experiments have shown that SL5's secondary structure forms a four-way junction that sequesters the genome's start codon within one of its helical stems (5–8, 10). Beyond SARS-CoV-2, the multi-way junction and the sequestration of the start codon in a stem are conserved features across most coronaviruses (6, 27, 28), despite large sequence divergence from SARS-CoV-2 (average sequence identity of 51% for the NCBI Reference Sequences for alpha- and betacoronaviruses; *SI Appendix*, Fig. S1). While conserved tertiary structural motifs would

Significance

The three-dimensional (3D) structures of viral RNAs are of interest to the study of viral pathogenesis and therapeutic design, but the 3D structures of viral RNAs remain poorly characterized. Here, we provide 3D structures of the SL5 domain from the majority of human-infecting coronaviruses. Comparison of SL5 structures suggests that conserved 3D elements may be important for SL5's as-of-yet-defined function. These conserved tertiary features support the relevance of SL5 for pan-coronavirus fitness and highlight routes towards understanding its molecular and virological roles and in developing SL5-based antivirals.

Author contributions: R.C.K., L.X., W.C., and R.D. designed research; R.C.K., L.X., I.N.Z., X.Z., R.H., S.L., and K.Z. performed research; R.C.K., L.X., I.N.Z., and R.H. contributed new reagents/analytic tools; R.C.K., G.N., K.Z., and R.D. analyzed data; and R.C.K., L.X., I.N.Z., W.C., and R.D. wrote the paper.

Reviewers: H.-W.W., Tsinghua University; and S.A.W., Johns Hopkins University.

Competing interest statement: Stanford University is filing patent applications based on concepts described in this paper.

Copyright © 2024 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

¹To whom correspondence may be addressed. Email: wahc@stanford.edu or rhiju@stanford.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2320493121/-/DCSupplemental>.

Published March 1, 2024.

be attractive targets for antiviral targeting and help pinpoint SL5 function, it is unknown whether SL5 forms a stable 3D structure in solution. Computational algorithms for RNA tertiary structure predict a wide range of 3D conformations (9, 29). Furthermore, SL5 stem lengths, internal loops, and sequences at the four-way junction are not conserved across coronaviruses (*SI Appendix, Tables S1 and S2*), leading to further uncertainty as to whether SL5 robustly forms a tertiary fold.

SL5 has been proposed to play a role in protein–RNA or RNA–RNA binding because of its conserved hexaloops, characterized by 5′-UUYUGU-3′ (Y = C, U) repeat loop motifs (27). This motif is conserved across most alpha- and betacoronaviruses, with the exception of betacoronaviruses in the *Embecovirus* subgenus, including the human-infecting HCoV-HKU1 and HCoV-OC43, which instead harbor a repeat loop motif elsewhere in their genomes (27). While the function of the UUYUGU hexalooop is unknown, it has been proposed to serve as a packaging signal in coronaviruses (27), and the 5′ proximal region has been confirmed to contain a packaging signal in one alphacoronavirus, the pig-infecting transmissible gastroenteritis coronavirus (30). The start codon’s occlusion in SL5’s secondary structure suggests that the folded form of SL5 does not enhance translation, and the deletion of sub-structures of SL5 increases its translation efficiency (31). However, the full role of SL5 structure in viral packaging or viral translation remains unclear and roles in other viral functions, such as viral genome replication, have not been ruled out.

To set a foundation for structure–function relationships, we set forth to characterize conserved structural elements among homologous constructs to distinguish species-specific and genus-specific features of SL5. Comparative structural biology studies have been widely pursued to assess the structures of multiple coronavirus spike proteins (32, 33). Such tertiary structural comparisons in RNA-only structures, however, have been limited. Among viral RNA structures, previous comparisons, conducted by NMR or X-ray crystallography, were limited to the comparison of two homologs (34–36). Structural comparisons for coronavirus RNA genomes have been limited to RNA secondary structure, including analysis of the 5′ UTR (26, 27) and the frameshift stimulation element (37), rather than comparisons of tertiary structure. Cryogenic electron microscopy (cryo-EM) offers the possibility of more routinely characterizing RNA elements across multiple homologs, particularly when integrated with biochemical secondary structure determination and automated computer modeling, such as in the Ribosome pipeline (38).

Herein, we conduct a comparative study of SL5’s tertiary fold across six coronaviruses. First, we report a tertiary structure of a 124-nt portion (40.0 kDa) of SARS-CoV-2 SL5, obtaining a 4.7 Å map of the domain. Contrary to the many possible tertiary conformations observed in de novo computational modeling (9, 29), we observe a well-resolved T-shaped conformation, where the larger stem-loops, SL5a and SL5b, base stack perpendicularly to the SL5-stem, with the short SL5c stem-loop jutting out from the T-shape. We then investigate 3D structural homology by resolving the structures of the SL5 domain from five additional coronaviruses. We find that all studied alpha- and betacoronaviruses share the same base stacking geometry, but the inter-helical angle is only shared within each genus. Additionally, within betacoronaviruses, the merbecovirus orthologs show tertiary features not observed in sarbecoviruses. Even though half of the SL5 domains characterized exhibit structural heterogeneity, every SL5 domain examined here populated a conformation in which the UUYUGU hexalooop sequences were displayed at opposing ends of a coaxial stack of conserved length. These structures and the analysis of 3D structural feature conservation suggest hypotheses for the function of the SL5 RNA element and may aid the rational design of therapeutics.

Results

SARS-CoV-2 SL5 Domain Secondary Structure. Our approach to SL5 structural characterization followed the Ribosome protocol (38), which integrates biochemical determination of secondary structure, tertiary fold determination with cryo-EM, and automated coordinate building in Rosetta. The secondary structure of SL5 was determined using multidimensional mutate-and-map chemical mapping as read out by sequencing (M2-seq) (39) with an updated “scarless” procedure in which, at the time of chemical modification, the sequence of interest does not include any flanking sequences. The secondary structure was found to contain a four-way junction without unpaired nucleotides, two UUYUGU hexaloops, and a GAAA tetraloop (Fig. 1 *A* and *B*). This secondary structure was also recovered from a “large-library” M2-seq approach, which used a synthesized library of mutants instead of error-prone PCR, and also matches previous *in vivo* and *in vitro* experimental secondary structure determinations (*SI Appendix, Fig. S2*) (5–8, 10). The secondary structures are consistent, with only minor differences in base-pairing at the terminal stem (residues 159 to 165 and 277 to 282) where our construct was excised out of the SARS-CoV-2 genome.

SARS-CoV-2 SL5 Exhibits a Stable 3D Tertiary Fold. Cryo-EM image reconstruction of the SARS-CoV-2 SL5 domain (residues 159 to 282, 124 nt, 40.0 kDa) shows a single, well-defined 3D structure resolved to 4.7 Å resolution (Fig. 1 *C–E* and *SI Appendix, Fig. S3*). Four helices extend from one junction, all with clear major and minor grooves. The approximate lengths of these helices align with the expected lengths of the stems in the M2-seq secondary structure, enabling the unambiguous identification of SL5c as the shortest stem and SL5b as the medium-length stem (Fig. 1 *E* and *SI Appendix, Table S3*). The junction is well resolved, revealing two pairs of coaxially stacked helices and a hole at the junction that clearly demarcates the backbone connectivity between helical pairs (Fig. 2*A*). This connectivity also enables the unambiguous assignment of SL5a and the SL5-stem into the map (Fig. 1*E*).

SARS-CoV-2 SL5 Domain 3D Structure. Guided by the 4.7 Å map and M2-seq secondary structure, integrative modeling was conducted using auto-DRRAFTER, which was specifically developed for medium to low-resolution RNA-only cryo-EM maps (38). Auto-DRRAFTER enumerates helical placements in the cryo-EM density and scores models by a combined biophysical and fit-to-map score. Acknowledging that in cryo-EM maps worse than 3.5 Å resolution, nucleotide bases cannot be precisely placed, we use the top 10 models, ranked by the auto-DRRAFTER scoring, as a representation of experimental uncertainty. Additionally, in this study, we identified multiple secondary structures from M2-seq, the literature (6), and predictions from EternaFold (40). All 3D modeling that converged across multiple computational runs and that fit well into the map was collected into a coordinate ensemble. The uncertainty in modeling is estimated by the modeling convergence, the mean pairwise root-mean-squared error (rmsd) as described previously (38). This resulted in an ensemble of 30 models with a convergence of 2.5 Å; slight differences in secondary structure are reflected in the ensemble but did not change the global tertiary fold (*SI Appendix, Fig. S4* and *Dataset S1*). Auto-DRRAFTER converged on a T-shaped conformation wherein the SL5-stem forms the “leg” of the T-shape, SL5a and SL5b stack end-to-end to form the perpendicular “arms” of the T-shape, and SL5c juts out of the plane of the T-shape at the junction. This automated modeling agrees with the global fold that we visually inferred from the map features alone.

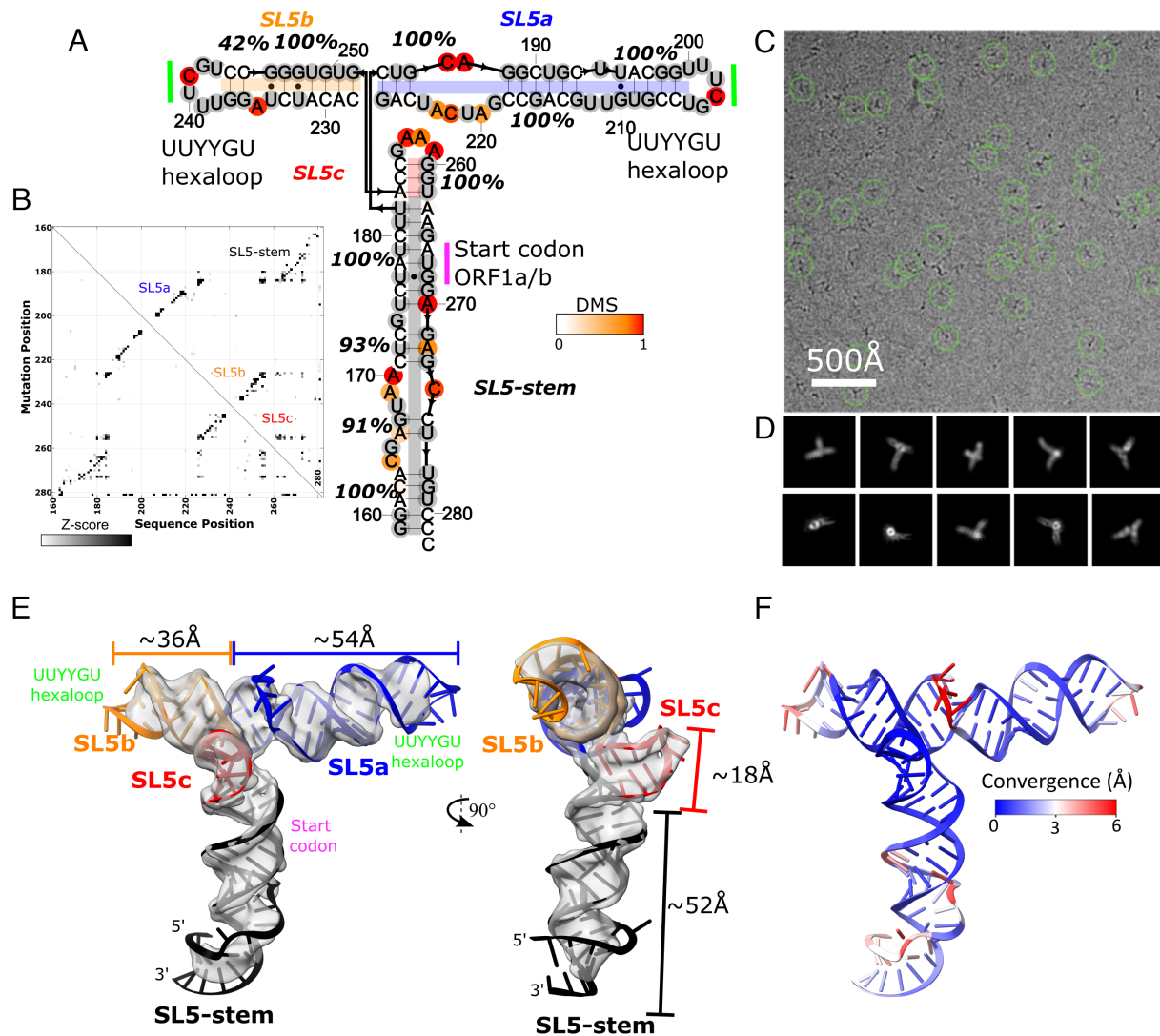


Fig. 1. The 3D global fold of SARS-CoV-2 SL5. (A) M2-seq-derived secondary structure of SARS-CoV-2 SL5. Stem confidence estimates (100 bootstraps) are given as percentage values and nucleotides are colored by DMS reactivity. (B) M2-seq Z-score plot, where the increases in reactivity across the molecules (x-axis) upon mutations (y-axis) are displayed in black. (C) Representative micrograph and (D) 2D class averages for the cryo-EM dataset of SARS-CoV-2 SL5. (E) The 4.7 Å cryo-EM map, displayed in gray, with a representative model. The model was obtained by using the M2-seq derived secondary structure and auto-DRRAFTER followed by refinement with ERRASER2. SL5 helices are colored in black (SL5-stem), blue (SL5a), orange (SL5b), and red (SL5c). The locations of the start codon (magenta) and UUYYGU hexaloops (lime) are also labeled. (F) A representative model is colored by per-residue convergence (mean pairwise r.m.s.d.) between the 30 models; blue areas are well converged, and red areas are divergent.

The UUYYGU hexaloops of SL5a and SL5b are positioned at the “hands” of the T-shape, positioned 84.2 ± 0.8 Å away from each other ($N = 30$, Fig. 1E and Table 1, and *SI Appendix, Fig. S5*). The two pairs of coaxially stacked helices, SL5a:SL5b and SL5-stem:SL5c have an inter-helical angle of $84.3 \pm 0.5^\circ$ but do not form any significant tertiary interactions ($N = 30$, Fig. 1E and Table 1, and *SI Appendix, Fig. S5*).

The auto-DRRAFTER models were further refined using ERRASER2 (version 2, available in Rosetta 3.10) (41) and evaluated for physical outliers and model-to-map fit (*Dataset S2*). The non-base-paired regions of the RNA converged the least during modeling, reflecting uncertainty in their tertiary structure (Fig. 1F). Such heterogeneity is supported by low map resolvability and Q-score (42) in these regions of the map (*SI Appendix, Fig. S4*). The four-way junction is well converged and has atomic resolvability above what is expected at 4.7 Å resolution (0.41 Q-score for junction atoms on average, with an expected 0.35 Q-score at that resolution). Refinement with ERRASER2 significantly improved the stereochemical quality of the models, while only marginally

changing model-to-map fit scores, indicating ERRASER2 was able to successfully refine these models (*SI Appendix, Fig. S6* and *Dataset S2*). Additionally, after refinement with ERRASER2, the models remained divergent in the regions with poor map resolvability, showing that the set of models continues to reflect experimental uncertainty post-refinement (*SI Appendix, Fig. S4* and *Dataset S1*). These trends hold true for modeling of all constructs in this study (*SI Appendix, Fig. S6*).

Substantiation of the SARS-CoV-2 SL5 Domain 3D Structure.

Next, we investigated an RNA segment, SL5-6 (residues 148 to 343, 196 nt, 63.1 kDa), which contains the full SARS-CoV-2 SL5 domain (residues 150 to 294) and its nearest neighboring domain, the SL6 domain (residues 302 to 343) (Fig. 3A). M2-seq experiments revealed that the SL5 secondary structure folds independently of the SL6 domain with a 7-nt linker region (residues 295 to 301) (*SI Appendix, Fig. S2*). In the cryo-EM map of SL5-6 (7.8 Å resolution, modeling convergence 4.4 Å *SI Appendix, Figs. S4* and *S7*), we resolved SL5, which retains

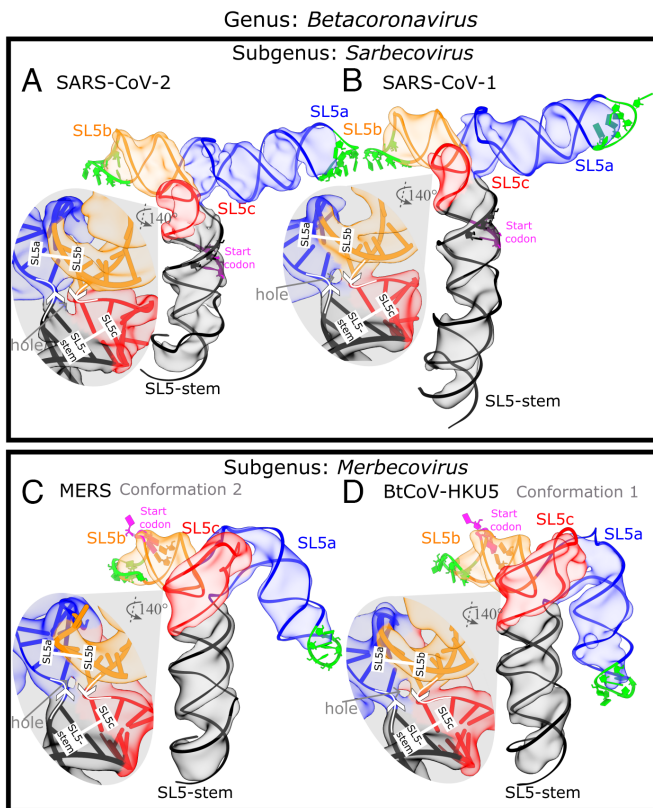


Fig. 2. The 3D global fold of SL5 in betacoronaviruses. The cryo-EM maps and representative model of the SL5 domain of two sarbecoviruses, (A) SARS-CoV-2 and (B) SARS-CoV-1, and two merbecoviruses, (C) MERS and (D) BtCoV-HKU5, are colored and labeled by stem. The nucleotides in the start codon are sequestered in a stem, and these base-paired nucleotides are displayed with the start codon in magenta. The nucleotides in the UUYUGU hexaloops are displayed in lime green. For each map, a zoom-in of the four-way junction cryo-EM map and model is displayed, showing the 5'-to-3' direction with white arrows, the coaxial stacking with white bars, and the junction hole with a gray arrow.

the previously observed T-shaped 3D fold (Fig. 1E) but did not observe density corresponding to SL6 (Fig. 3 B and E and *SI Appendix, Fig. S7*). To test the stem assignments in the SL5-6 structure and verify the absence of SL6, we designed extensions in the SL5b and SL5c stems (both 204 nt, 65.7 kDa) so that observed changes in the cryo-EM map would tag the corresponding stems,

as previously done (12, 43–45). The tertiary structures of the extension constructs (SL5b extension: 7.4 Å resolution, modeling convergence: 2.7 Å; SL5c extension: 9.0 Å resolution, modeling convergence: 2.8 Å *SI Appendix, Figs. S4 and S7*) exhibited new densities in the extended stems, substantiating the stem assignments in our SL5-6 and SL5 models (Fig. 4 C, D, F, and G). Though these maps (deposited to EMDDB) were modeled successfully by auto-DRRAFTER, the corresponding models were not deposited to the PDB due to the partial resolvability of the construct; the maps only resolve part of the RNA construct, which may have led to the distortions at the apical ends of stems in the model (*SI Appendix, Fig. S4*).

We hypothesized that SL6 was not well resolved due to flexibility in the natural linker sequence connecting SL5 and SL6. This hypothesis is supported by an additional density consistent with SL6 appearing adjacent to the SL5-stem when SL5a, SL5b, and SL5c are removed, either by subtracting their density from particle images or from imaging an RNA construct (129 nt, 41.4 kDa) without these stems (*SI Appendix, Fig. S8*). Furthermore, unlike in the map containing the full SL5, in the map containing the SL5-stem with SL6, the helical grooves of the SL5-stem are not resolved, consistent with averaging that would result from SL6 moving around the SL5-stem (*SI Appendix, Fig. S8*). The natural extension of SL5 by SL6, along with the designed extensions, preserve SL5's distinct T-shaped fold, which suggests that the tertiary structures of SL5 and SL6 fold independently of each other.

3D Structure of the SL5 Domain in Betacoronaviruses. Based on our cryo-EM analysis, the SL5 domain of SARS-CoV-2 has a defined tertiary fold, as opposed to a highly flexible ensemble. The coronavirus genome may have evolved this sequence to have a specific arrangement of SL5's four stems to serve a functional role. However, this arrangement could also be a biophysical coincidence and not be a result of natural selection. While the secondary structure of the SL5 domain is conserved in most coronaviruses, it is unknown whether its tertiary structure would also be conserved, given the region's low sequence conservation (pairwise sequence identity mean of 54.3% and minimum of 46.9% for the six sequences studied here, *SI Appendix, Supplemental Note S1*) and variation in stem lengths. 3D conservation of the arrangement of these stems across coronaviruses would further support the importance of this feature for SL5 function.

Table 1. 3D features of the SL5 domain of coronaviruses

Genus	Subgenus	Species	Distance between UUYUGU hexaloops (Å) [*]	Inter-helical angle (°) [†]	
<i>Betacoronaviridae</i>	<i>Sarbecovirus</i>	SARS-CoV-2	84.2 ± 0.8	84.3 ± 0.5	
		SARS-CoV-1	82 ± 1	86.8 ± 0.5	
	<i>Merbecovirus</i>	MERS	80 ± 2	82 ± 2	
		BtCoV-HKU5	80 ± 4	86 ± 2	
<i>Alphacoronaviridae</i>	<i>Duvinacovirus</i>	HCoV-229E	92.4 ± 0.9	-121.3 ± 0.2	
			SL5b ↔ SL5c	94.3 ± 0.8	
			SL5a ↔ SL5c	49.1 ± 0.8	
<i>Setracovirus</i>	HCoV-NL63		95, 90 [‡]	-120 ^{‡,§}	
			SL5b ↔ SL5c	85, 65 [‡]	
			SL5a ↔ SL5c	45, 50 [‡]	

^{*}Distance is between the apical loops of SL5a and SL5b unless otherwise specified. Distance is defined as the distance between the centroid of the C1' atoms of each loop. Error is SD of the distance across the refined auto-DRRAFTER models.

[†]The angle is the angle between the SL5-stem:SL5c coaxial stack and SL5a:SL5b coaxial stack. The parallel configuration is defined as 0° and the anti-parallel as 180°. The positive rotation is defined as rotating the SL5a:SL5b stack, positioned in the background, clockwise as in the orientations in Fig. 2. See *SI Appendix, Fig. S5* for a diagram explaining the angles. Error is the SD of the angle across the refined auto-DRRAFTER models.

[‡]Cryo-EM maps were not modeled, so distance and angle were estimated from the map alone.

[§]Only conformation 1 is considered, as conformation 2 does not have the listed coaxial stacks.

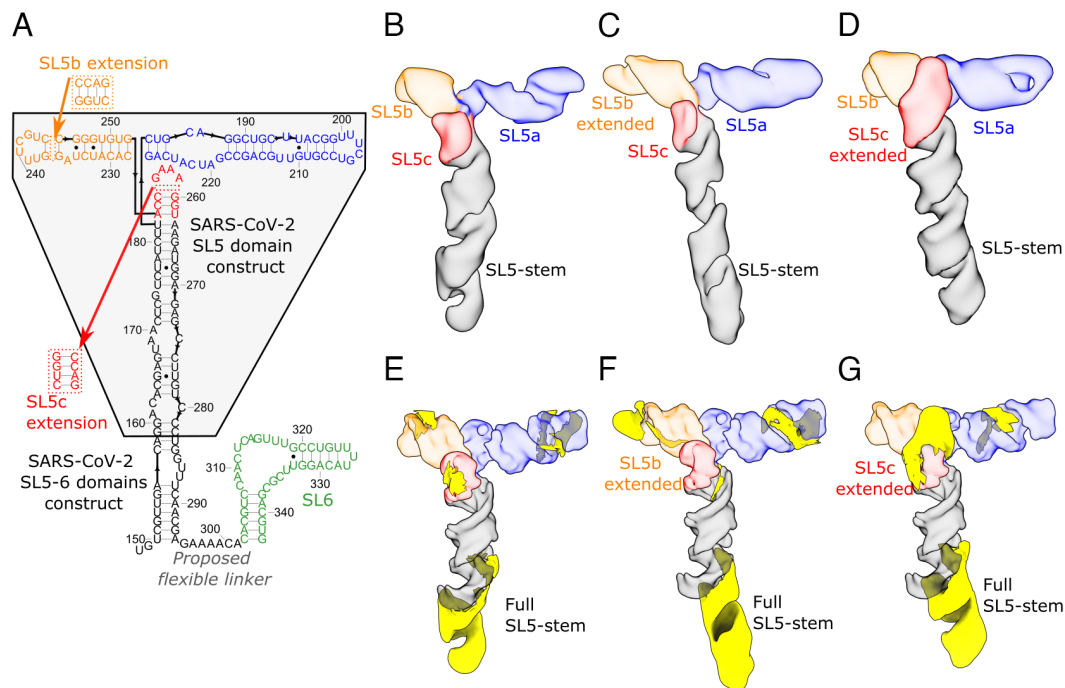


Fig. 3. Substantiation of SARS-CoV-2 SL5 domain 3D structure using extension constructs. (A) The secondary structure, derived from M2-seq, of the SL5-6 domains of SARS-CoV-2 is depicted and colored in black (SL5-stem), blue (SL5a), orange (SL5b), red (SL5c), and green (SL6). The original construct used in this study for SARS-CoV-2 SL5 (Fig. 1) is highlighted in the gray box. Relative to this construct, all SL5-6 constructs are extended to include the full SL5-stem and an additional stem-loop, SL6. In addition, the location of the four base-pair extensions to SL5b and SL5c are depicted. The cryo-EM maps of (B) SL5-6, (C) SL5-6 with SL5b extended, and (D) SL5-6 with SL5c extended are displayed, colored, and labeled by stem. Extensions are highlighted in yellow after masking out the density of the original SL5 construct for (E) SL5-6, (F) SL5-6 with SL5b extended, and (G) SL5-6 with SL5c extended.

We therefore carried out cryo-EM to resolve the SL5 domain of other coronaviruses, with a focus on human-infecting coronaviruses, for structural comparison with the SL5 domain of SARS-CoV-2. Betacoronaviruses contain five human-infecting coronaviruses, SARS-CoV-2, SARS-CoV-1, MERS, HCoV-OC43, and HCoV-HKU1. Among these viruses, HCoV-OC43, and HCoV-HKU1 are members of the *Embecovirus* subgenus, which was previously found to have replaced UUYUGU hexaloops in SL5 with repetitive loop motifs elsewhere in the genome (27), and so our studies focused on SARS-CoV-1 and MERS.

First, we examined the SL5 domain from SARS-CoV-1 (residues 151 to 291, 143 nt, 46.1 kDa), which has high sequence similarity with SARS-CoV-2 (85.9% sequence identity, *SI Appendix, Supplemental Note S1*) and also belongs to the *Sarbecovirus* subgenus. We found that SL5 of SARS-CoV-1 (7.0 Å resolution, 2.6 Å modeling convergence, *SI Appendix, Figs. S4 and S9*) adopts the same T-shaped fold and junction geometry as that of SARS-CoV-2, with an inter-helical angle of $86.8 \pm 0.5^\circ$ and distance between UUYUGU hexaloops of 82 ± 1 Å ($N = 20$, Fig. 2B and Table 1 and *SI Appendix, Fig. S4*).

Expanding the cryo-EM analysis to more distant coronavirus relatives, we examined the orthologous SL5 domain from MERS (residues 206 to 338, 135 nt, 43.7 kDa), which belongs to the different *Merbecovirus* subgenus of betacoronaviruses. From the secondary structures obtained from large-library M2-seq and the literature (6), we already noted a difference: while the UUYUGU hexaloops are still found on SL5a and SL5b, the SL5c stem from MERS is significantly longer than SL5c from sarbecoviruses (*SI Appendix, Fig. S2*). Interestingly, our MERS SL5 cryo-EM analysis showed three conformations (6.9, 6.4, 6.4 Å resolution, 3.5, 3.2, 3.4 Å modeling convergence, *SI Appendix, Figs. S9 and S10*) and had the same conformation seen in sarbecovirus orthologs: helical stacking, junction geometry, and inter-helical angle matching

within experimental error (Fig. 2C and Table 1 and *SI Appendix, Fig. S10*).

To investigate the conservation of the SL5 fold further, we examined an additional merbecovirus ortholog, the SL5 domain of BtCoV-HKU5 (residues 188 to 320, 135 nt, 43.7 kDa). BtCoV-HKU5 SL5 has a 77.0% sequence identity to MERS SL5 (*SI Appendix, Supplemental Note S1*) and large-library M2-seq and the literature (6) show the secondary structure is very similar to MERS SL5 (*SI Appendix, Fig. S2*). The cryo-EM structure analysis of BtCoV-HKU5 SL5 resolved four conformations (5.9, 6.4, 8.0, 7.3 Å resolution, 3.0, 3.0, 5.2, 3.0 Å modeling convergence, *SI Appendix, Figs. S10 and S11*) and again shares the same helical stacking and junction geometry with the other betacoronaviruses studied, with an inter-helical angle of $86 \pm 2^\circ$ (Fig. 2D and Table 1 and *SI Appendix, Fig. S5*). While auto-DRRAFTER was able to model all maps consistently, we deposited all these maps to EMDB but did not deposit the coordinates for conformation 3 in the PDB because the helical grooves were insufficiently resolved (*SI Appendix, Fig. S10*). As with the other three betacoronavirus domains imaged, the four-way junction is well resolved in the BtCoV-HKU5 maps, revealing a clear hole between the helical strands that separates perpendicular, coaxially stacked stems (Fig. 2D).

While the four-way junction geometry is conserved between the sarbecovirus and merbecovirus orthologs, the merbecoviruses have a distinct ensemble of 3D folds that is conserved between MERS SL5 and BtCoV-HKU5 SL5. In particular, a flexible bend is observed emanating from the SL5a internal loop that allows the SL5a arm to swing in a hinge-like motion (Fig. 4A and B). To model this conformational heterogeneity, the particles were classified into discrete classes to resolve cryo-EM maps that could individually be modeled (Fig. 4C and D). Among these conformations, the junction geometry was conserved, with inter-helical angles occupying a narrow range of 81 to 84° and 84 to 88° for MERS

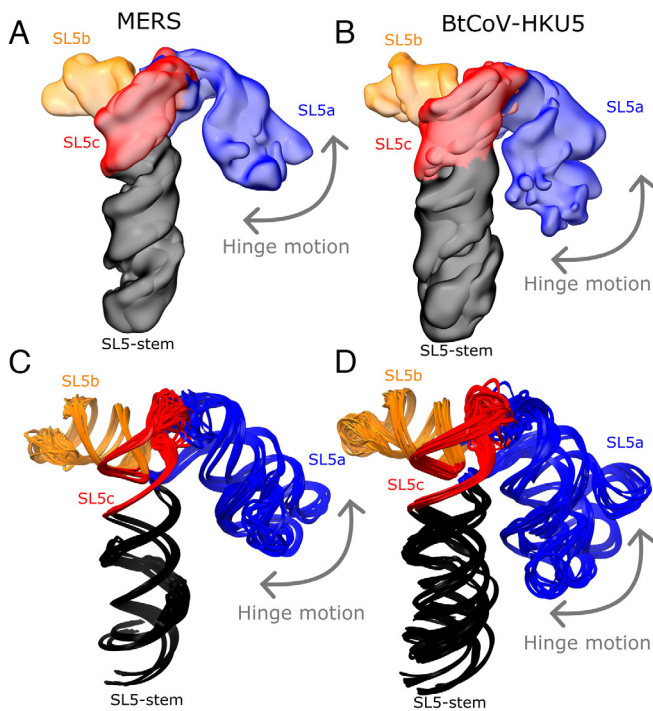


Fig. 4. The hinge motion in the SL5a stem of merbecoviruses. The cryo-EM maps obtained after discrete classification of particles are overlaid for the (A) MERS and (B) BtCoV-HKU5 SL5 domains. The models derived from these maps using auto-DRRAFTER and ERRASER2 are overlaid for the (C) MERS and (D) BtCoV-HKU5 SL5 domains. All constructs are colored by stem, and the hinge motion is labeled with an arrow.

and BtCoV-HKU5 SL5, respectively (SI Appendix, Fig. S5). Despite the junction geometry conservation, the relative locations of the hands of the SL5a and SL5b arms are variable due to the hinge in the SL5a internal loop, increasing the range of distance between UUYUGU hexaloops to 74 to 85 Å and 71 to 85 Å for MERS and BtCoV-HKU5 SL5, respectively (SI Appendix, Fig. S5).

In addition, the merbecovirus orthologs display an unexpected tertiary interaction between the SL5a internal loop and SL5c apical loop (Fig. 5 C and E), whereas the sarbecovirus SL5c stem is too short to form this tertiary interaction (Fig. 5 A and B). Due to resolution limitations, the SL5a–SL5c interaction of merbecoviruses cannot be modeled with atomic precision. This uncertainty in the SL5a–SL5c interaction is reflected in the ensemble of models produced by auto-DRRAFTER (Fig. 5 D and F). The models, across all conformations and both merbecovirus orthologs, do converge in identifying the same interacting regions, namely the asymmetric internal loop of SL5a, 5'-AAUU-3' and the apical loop of SL5c, 5'-AAGGUGC-3' (MERS: residues 264 to 267 and 397 to 313, respectively; BtCoV-HKU5: residues 246 to 249 and 289 to 295, respectively Fig. 5 G and H).

3D Structural Comparison of the SL5 Domain across Alpha- and Betacoronaviruses. We next looked to alphacoronaviruses to explore the 3D structure of SL5 in a different genus. We selected the SL5 domains from the remaining two human-infecting coronaviruses, HCoV-229E (residues 153 to 292, 140 nt, 45.1 kDa) and HCoV-NL63 (residues 138 to 295, 160 nt, 51.4 kDa) from the *Duvinacovirus* and *Setracovirus* subgenera, respectively, for investigation. While an experimental secondary structure for the SL5 domain in HCoV-NL63 was previously identified (6), the secondary structure for HCoV-229E was deduced by exhaustively modeling a set of published (26, 27), predicted (40, 46), and manually curated secondary structures into the cryo-EM map

(SI Appendix, Fig. S12). Only one secondary structure for HCoV-229E resulted in converged auto-DRRAFTER modeling that agreed with the cryo-EM map (SI Appendix, Fig. S13). Beyond containing four helical stems, the secondary structures of these alphacoronaviruses differ from those of the previously examined betacoronaviruses. These alphacoronaviruses have three UUYUGU hexaloops, as opposed to two, and the four-way junctions contain unpaired nucleotides. Hence, we sought to investigate which 3D structural features, if any, were conserved within human-infecting alphacoronaviruses and between alpha- and betacoronaviruses.

HCoV-229E SL5 (6.5 Å resolution, 2.3 Å modeling convergence, SI Appendix, Figs. S9 and S13) and HCoV-NL63 SL5 (8.0, 8.4 Å resolution, not modeled, SI Appendix, Fig. S14) form X-shaped folds. Both alphacoronavirus SL5s adopt the same helical stacking as the betacoronavirus domains, with the SL5-stem:SL5c and SL5a:SL5b coaxially stacking (Fig. 6). The junction is well resolved in HCoV-229E, with a visible hole separating the stems (SI Appendix, Fig. S13). For HCoV-NL63, however, the cryo-EM data were classified into two maps with distinct conformations (SI Appendix, Figs. S13 and S14). These maps did not achieve sufficient resolution to view the major or minor grooves of helices, and thus, coordinates were not modeled into the maps. Nevertheless, the disparate lengths of each stem in HCoV-NL63 SL5 allow for unambiguous stem assignment to assess junction geometry (SI Appendix, Fig. S13). One conformation reveals a distinct stacking pattern at the junction, in which the SL5-stem stacks coaxially with SL5a while SL5b stacks with SL5c. The other conformation is homologous with the HCoV-229E global fold that coaxially stacks SL5a and SL5b, as seen with all the other coronaviruses studied here.

The inter-helical angle is similar among the alphacoronavirus SL5s, $-121.3 \pm 0.2^\circ$ ($N = 10$) and -120° (estimated) for HCoV-229E and HCoV-NL63 respectively, despite being quite different in sequence (59.1% sequence identity, SI Appendix, Supplemental Note S1, Fig. 3 and Table 1). This inter-helical angle for X-shaped alphacoronavirus SL5s is distinct from the near-perpendicular angle formed by the T-shaped betacoronaviruses SL5s (81 to 88°) (Fig. 6). Despite this difference, the UUYUGU hexaloops are also positioned a similar distance apart of 92 and 95 Å for HCoV-229E SL5 and HCoV-NL63 SL5 conformation 1, respectively (Table 1). Surprisingly, despite a different coaxial stacking pattern, HCoV-NL63 conformation 2 has a similar estimated SL5a:SL5b UUYUGU hexalooop distance of 90 Å. These distances are ~ 10 Å longer than the distances, 82 to 84 Å, observed in the betacoronavirus SL5 domains. While not identical distances, the observed range is narrower than the ranges of distance predicted by de novo 3D structure prediction algorithms in CASP15 for SARS-CoV-2 and BtCoV-HKU5 SL5 (29) (SI Appendix, Fig. S5).

Discussion

We have presented structural characterization of the SL5 domain across six coronaviruses. This study was enabled by the increasing throughput of 3D RNA structural characterization, made possible by single particle cryo-EM integrated with biochemical secondary structure mapping, automated computer modeling, and structure validation (38, 42). Cryo-EM offers the opportunity to increase the knowledge base of RNA 3D global folds, particularly through the ability to study RNA homologs, as carried out here, revealing similarities and differences that may be relevant to function. While cryo-EM of these small RNA samples (40.0 to 65.7 kDa in size) was limited in resolution, for all but one sample, we were able to achieve sufficient resolution to

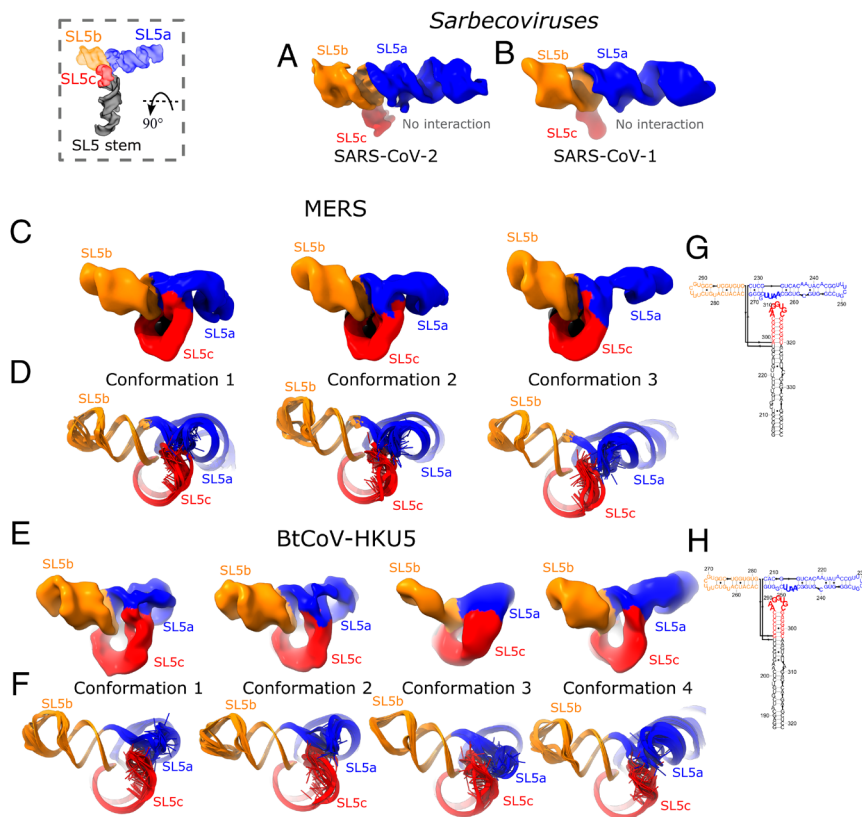


Fig. 5. The tertiary interaction in the SL5 domains of merbecoviruses. The proposed tertiary interaction is displayed by viewing the top of the molecules with the SL5-stem at the back, relative rotation is shown in the top left of the figure. (A and B) The SL5 domains from members of the *Sarbecovirus* subgenus do not have densities connecting SL5c (red) and SL5a (blue), indicating there is no tertiary interaction between these stem-loops. In SL5 domains from the *Merbecovirus* subgenus, (C and D) MERS and (E and F) BtCoV-HKU5, all cryo-EM maps and models are displayed and colored by stem. Cryo-EM maps (C and E) show that a density connecting the apical loop of SL5c to the first internal loop of SL5a was resolved. The cryo-EM maps were insufficiently resolved for modeling to converge on the atomic level details of this junction (D and F), but the same set of residues are consistently interacting. These interacting nucleotides are bolded in the secondary structures of (G) MERS and (H) BtCoV-HKU5.

resolve major and minor grooves and even resolve the hole in the four-way junction. This resolution enabled the unambiguous identification of stem positions and hence the RNA's global fold (Figs. 1, 2, and 6). This demonstrates the utility of cryo-EM to resolve the 3D folds of some viral RNA elements that may exhibit flexibility (Fig. 4). We further tested our structural models, demonstrating that the SL5 domain folds independently of the closest downstream stem-loop, SL6, and confirmed our stem placements using extension constructs (Fig. 6). Although atomic level detail cannot be ascertained from the structures, we have deposited multiple models in the Protein DataBank to record this experimental uncertainty.

The SL5 domains of all human-infecting coronaviruses that contain the UUYUGU hexaloops, along with the bat-infecting BtCoV-HKU5 SL5, fold into a limited number of stable conformations, which we have resolved using cryo-EM. We found that the SL5 domain of SARS-CoV-2 folds into a T-shaped structure, modeled using a 4.7-Å cryo-EM map, in which the SL5-stem and SL5c form a continuous, long, coaxial stack that lies perpendicular to a second continuous, long, coaxial stack formed by SL5a and SL5b. This stacking pattern is conserved across all of the imaged SL5 domains while the junction geometry and inter-helical angles of 81 to 88° for betacoronaviruses and approximately -120° for alphacoronaviruses are conserved within each genus. Although the junction angle is genus-specific, all coronaviruses studied display an experimentally resolved conformation that places a pair of UUYUGU hexaloops a distance of 82 to 92 Å apart at opposing ends of an SL5a:SL5b coaxial stack.

The *Merbecovirus* subgenus of betacoronaviruses has an additional subgenus-specific structural feature: an interaction between

the SL5a internal loop and SL5c apical loop (Fig. 5). Signatures for this interaction have not been observed in chemical mapping studies, which are frequently not sensitive to tertiary interactions (47, 48). The atomic structure of this feature was not resolved, and future work could improve the local resolution of this interaction by taking into account inherent flexibility in other regions of the RNA (49, 50). This interaction could help stabilize the SL5 stacking pattern and junction orientation, relative to other global stacking patterns and inter-helical rotations that would position the SL5a internal loop and SL5c apical loop apart.

The structurally conserved features of SL5 across coronavirus genera suggest potential functional roles for SL5. First, the SL5 domain sequesters the start codon in a stem, but this sequence must be exposed by unfolding SL5 to initiate translation. Thus, the SL5 element may act as a switch, enforcing exclusivity between viral translation and an as-yet-unknown function corresponding to SL5's folded structure, such as viral replication or viral packaging, that should not occur at the same time as viral translation. Work on the structure of SL5 bound to translation initiation machinery may elucidate the nature of the conformational change required for translation.

Second, SL5 contains two or three UUYUGU hexaloops in the selected betacoronavirus and alphacoronavirus domains, respectively. These cryo-EM structures reveal that the distances between hexaloops are within a narrow range, 82 to 92 Å, and are placed at opposing ends of a long, continuous coaxial stack. This conservation suggests that SL5 may position the hexaloops for a functional reason, but this hypothesis remains untested. As one possibility, the viral genome must be selectively packaged, compared to host RNA,

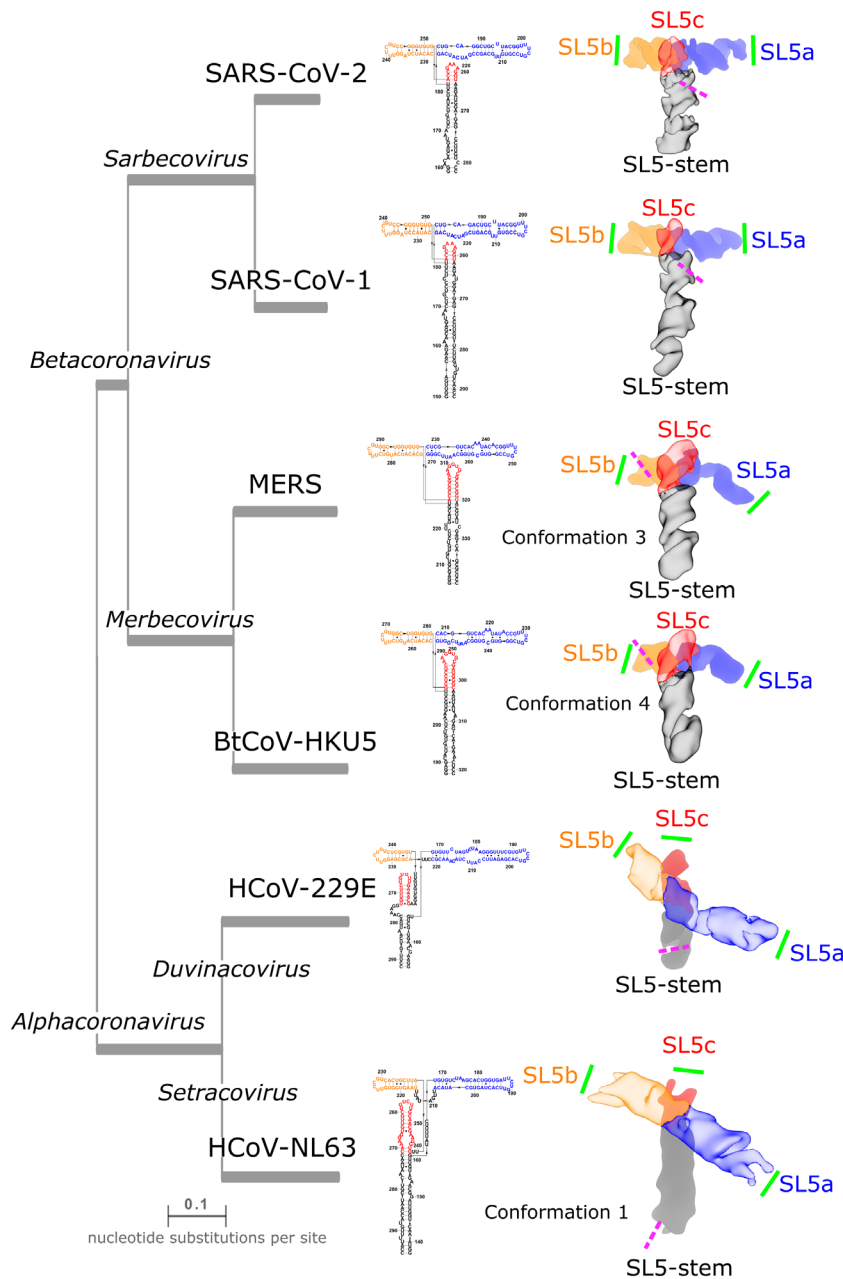


Fig. 6. Similarities and differences in the tertiary folds of SL5 in alpha- and betacoronaviruses. To compare junction geometries between the SL5 domain of betacoronaviruses and alphacoronaviruses, the cryo-EM maps of the SL5 domains are displayed with the junction perpendicular to the text and with the stem pointed downward as a reference. Additionally, SL5b is pointed left and SL5a right. The maps are colored by domain with the foreground, SL5-stem:SL5c for betacoronaviruses SL5ba:SL5b for alphacoronaviruses, made transparent to enable the view of the stems below. The various species are positioned on a phylogenetic tree with branch length proportional to evolutionary distance and secondary structures, arranged to correspond to the cryo-EM structure, are displayed. For BtCoV-HKU5, MERS, and HCoV-NL63 SL5, the conformation that is closest to other SL5 domains is displayed. The pink dashed lines indicate the position of the start codon; note that for HCoV-NL63, the SL5 construct imaged was truncated on the 3' end directly before the start codon. The green solid lines indicate the positions of UUYUGU hexaloops.

in virions. While UUYUGU motifs will recur throughout host RNAs by chance, the stereotyped placement of two such sequences as apical loops on opposing sides of a coaxial stack are less likely to occur in host RNAs. The two loops could therefore be selectively recognized by oligomers of viral or host proteins with RNA-binding domains (51). While, in principle, structures other than four-way junctions can produce similar coaxial stacks, the SL5 four-way junction provides a natural solution. For example, three-way junctions can provide similar positionings of UUYUGU motifs at opposite ends of a co-axial stack, but they can form a larger number of alternative stacking patterns than four-way junctions, which would give rise to competing, potentially non-functional geometries (52).

In addition to resolving dominant structures of the SL5 elements, since we imaged RNA in vitreous ice, we were able to resolve a “cryoensemble” of structures. We observed a flexible hinge of SL5a in both merbecovirus SL5 domains and an alternative stacking pattern at the four-way junction of the HCoV-NL63 SL5 domain. Additionally, we observed indications of other modes of flexibility that we did not model—for example, the SL5-stem in SARS-CoV-1 (*SI Appendix, Fig. S9*). Despite this, after taking experimental uncertainty into account, we conclude that all betacoronavirus SL5s imaged exhibit interhelical angles within a narrow range, despite modeled and unmodeled flexibility. This interhelical angle range is significantly different from that observed in the alphacoronavirus SL5s.

We observed some limitations in current cryo-EM data analysis procedures. For example, we hypothesize that the 7-nt flexible linker between SL5 and SL6 left SL6 (13.3 kDa) unresolved. Also heterogeneity likely limited the angular assignment accuracy and resolution of cryo-EM maps, particularly in the case of the HCoV-NL63 SL5. These shortcomings highlight the need for further methods to be developed and tested on RNA constructs, which may reveal additional, unobserved heterogeneity, especially for small, helical structures with continuous hinge motions. Additionally, while imaging RNA in vitreous ice is a step toward achieving more near-native conditions, the effects of excising these RNA elements from genomic and cellular contexts, as well as the effects of the grid environment and vitrification on RNA structure and heterogeneity are unknown. Complementary, lower-resolution experimental techniques such as single-molecule FRET or solution X-ray scattering paired with molecular dynamics could be used to further understand biases of the different methods and more quantitatively assess the relative populations of RNA species in solution (53, 54).

Despite these caveats, the ability to solve the cryoensemble for such small RNA molecules, although likely not representing the full, biologically relevant ensemble, was important for understanding the conservation of the 3D fold in the coronaviruses' SL5 domain. It is possible that the crystallized structures of these RNAs would not have revealed the same conservation as evident when we analyze the cryoensemble—for example, the SL5 domain from HCoV-NL63 may crystallize as conformation 2, the conformation with alternative base-stacking. Alternative conformations may have distinct functional roles, and future work could aim to test significance of alternative conformations. An “ensemble view” of RNA molecules enhances structure–function interpretations and may be more readily brought to bear in RNA systems in the future through cryo-EM (55).

Finally, the analysis of conserved structural features of the SL5 domain suggests strategies for the structure-guided design of pan-coronavirus therapeutics. In particular, there may be druggable pockets at the four-way junction, conserved among the betacoronaviruses studied here, or at the SL5a–SL5c tertiary interaction in MERS, the human-infecting coronavirus with the highest fatality rate. These pockets could be the targets for small molecules such as ribonuclease-targeting chimeras (RIBOTACs) (20). Alternatively, targeting two regions could improve the specificity of a therapeutic. For example, antisense oligonucleotides that target both the start codon and the four-way junction region of SL5 would serve the dual purpose of slowing viral translation while also preventing formation of the SL5 tertiary structure, which appears important for a distinct viral function. Different classes of therapeutics could also take advantage of the stereotyped positioning of the conserved UUYUGU hexaloops. For example, circularized or chemically modified RNAs could present UUYUGU hexaloops positioned 82 to 92 Å apart and thereby compete for the binding of proteins with the viral genomic RNA. Additionally, the catalog of SL5 structures may enable faster response for emerging threats by enabling the design of therapeutics against orthologous structures from any of the viruses resolved herein, which represent most known human-infecting coronaviruses. Finally, the models of the cryoensemble could present an opportunity for structure-guided drug design, enabling the targeting of one conformation over the others with the aim of trapping the RNA in a non-functional conformation. These efforts will likely require complementary contributions from X-ray crystallography and NMR to resolve higher resolution details important for designing and refining structure-guided small molecule therapeutics.

Materials and Methods

“Scarless” 2D Chemical Mapping. 2D chemical mapping was performed using an optimized M2-seq pipeline (39), which uses mutational sequencing-based inference of dimethyl sulfide (DMS) modifications on a library of folded RNA that contains purposeful random sequence variations to better infer stems. The scarless protocol was modified to remove primer binding sequences from the RNA to prevent unwanted secondary structure interference during DMS modification. This modification was achieved by appending removable primer sequences to the 3' end of the DNA encoding the region of interest, which is used for error-prone PCR (epPCR) and cleaved off prior to *in vitro* transcription. After DMS modification, sequencing libraries were made using two ligation steps on ssRNA and ssDNA followed by a primer-biased PCR. See *SI Appendix, Supplemental Note 3* for detailed methods.

“Large-library” 2D Chemical Mapping. To accelerate mutate-and-map characterization of secondary structures of multiple orthologs, we explored the use of oligonucleotide libraries that encoded SL5 domains as well as all of their single mutants, with 3' barcode hairpins to allow unambiguous deconvolution of the mutant profiles. The libraries then underwent a similar procedure to above, *in vitro* transcription, chemical modification, library preparation, and sequencing. See *SI Appendix, Supplemental Note 3* for detailed methods.

Secondary Structure Modeling. Chemical mapping profiles acquired in the mutate-and-map experiments above were analyzed with Biers (<https://ribokit.github.io/Biers/>) to generate normalized 1D DMS profiles and 2D Z-scores. Biers was then used to create secondary structure predictions guided by the 1D DMS profiles and 2D Z-scores using ShapeKnots, with 100 bootstrapping iterations to estimate stem confidence values. The secondary structure with the 1D DMS profile was depicted using RiboDraw (56). The raw data and Z-score plots were visualized using custom scripts. All scripts can be found in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D).

Sample Preparation for Cryo-EM. Primers to assemble the sequences (sequence of interest with a T7 promoter) were designed for PCR assembly using Primerize (listed in *Dataset S3*) (57), ordered from Integrated DNA Technologies, assembled into full-length double-stranded DNA by PCR assembly following the Primerize protocol using Phusion polymerase (*in-house*), and purified with QIAquick PCR Purification Kit (QIAGEN, #28104). RNA was synthesized by *in vitro* transcription (TranscriptAid T7 High Yield Transcription Kit, Thermo Scientific #K0441), then purified by column purification (RNA Clean & Concentrator Kits, Zymo Research #R1017) and by denaturing PAGE gel extraction (ZR small-RNA PAGE Recovery Kit, Zymo Research #R1070). RNA concentration was measured using a NanoDrop. Purified RNA was refolded prior to sample vitrification as follows: Purified RNA was diluted to a target concentration of 20 to 30 μM in 50 mM Na-HEPES, pH 8.0, denatured at 90 °C for 3 min, then cooled at room temperature for 10 min. RNA was incubated with 10 mM MgCl₂ at 50 °C for 20 min and then cooled at room temperature for 10 min. Then, 3 μL of refolded RNA was frozen by Vitrobot Mark IV (2.5 to 4 s blot time, 1 to 5 s wait time) onto Quantifoil R 2/1 grids or Quantifoil R 1.2/1.3 grids following glow discharge (30 s glow, 15 s hold). Refer to *SI Appendix, Table S4* for details on target RNA concentrations, grid type, and sample freezing conditions for each sample.

Cryo-EM Data Acquisition. Cryo-EM data were collected on a 300 kV Titan Krios G3i. For all data collections, EPU was used for screening, beam alignments, and automated collection. Digital micrograph and Sherpa were used for energy filter alignments for the BioQuantum and Selectris energy filter, respectively. For all data collections, a 100-μm objective aperture was inserted. Refer to *SI Appendix, Table S5* for exact values for detector, energy filter, energy filter slit size, nominal magnification, pixel size, total dose, dose per frame, frame duration, exposure time, and number of acquired micrographs for each sample.

Cryo-EM Data Processing. Single-particle image processing and 3D reconstruction was performed using CryoSPARC 3.2.0 (58). Patch motion-correction and patch CTF-estimation were used in pre-processing. Information regarding the data processing for each dataset can be found in *SI Appendix, Table S6* and the pipelines to process each dataset can be seen in *SI Appendix, Figs. S3, S7, S9, S11, and S14 and Supplemental Note S2*. General strategies included one to three iterations of 2D

classification with a larger box size to remove unfolded and aggregated RNA particles, which appear as elongated classes only when a larger box size is used; and 3D heterogeneous refinement using a spherical class to remove noise particles common when analyzing small particles which have a low signal-to-noise ratio.

Modeling Cryo-EM Maps. All cryo-EM maps were modeled using auto-DRRAFTER, except for maps where major grooves were not resolved: the HCoV-NL63 SL5 and SARS-CoV-2 SL5-6 with SL6 extended and SL5a, SL5b, and SL5c removed. All secondary structures identified were used in separate auto-DRRAFTER runs, resulting in multiple ensembles of models for each map. Notably, for the SARS-CoV-2 SL5-6 domains, SARS-CoV-2 SL5-6 domains with SL5b extended, and SARS-CoV-2 SL5-6 domains with SL5c extended, the sequences and secondary structures were cut off at the bottom of the SL5-stem. For all other constructs, the full sequences were used for modeling. For all modeling, the sharpened map was used, except for the SARS-CoV-2 SL5-6 domains; however, given the low-resolution of each map, the sharpened and unsharpened maps are not expected to result in different modeling results. Using auto-DRRAFTER, nodes were fitted into the map after it was low pass filtered at 20 Å using a map threshold specified in the legend of *SI Appendix, Figs. S4, S10, and S13*. The helical placements were exhaustively searched by initially placing the ends of the four stems in an "end-node," as displayed in *SI Appendix, Figs. S4, S10, and S13*. This node was consistent for each secondary structure modeled. Rounds of 5,000 decoys for each initial stem placement were run until convergence, defined as less than 10 Å mean pairwise r.m.s.d. between the top 10 models. After these initial runs, two final rounds were run, also creating 5,000 decoys, to obtain a final set of top 10 models. The exact commands can be found in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D). The 10 final auto-DRRAFTER models were refined using ERRASER2 (<https://new.rosettacommons.org/docs/latest/ERRASER2>) with the following command in Rosetta:

```
eraser2
-s $PDB
-edensity:mapfile $MAP
-edensity::mapreso $RESOLUTION
-score:weights stepwise/rna/rna_res_level_energy7beta.wts
-set_weights elec_dens_fast 10.0 cart_bonded 5.0 linear_chainbreak
10.0 chainbreak 10.0 fa_rep 1.5 fa_intra_rep 0.5 rna_torsion 10 suite-
ness_bonus 5 rna_sugar_close 10
-rmsd_screen 3.0
-mute core.scoring.CartesianBondedEnergy core.scoring.electron_dens-
ity.xray_scattering
-rounds 3
-fasta $FASTA
-cryoem_scatterers
```

The 10 models were then combined into a single PDB file and `pdb_extract` (<https://pdb-extract.wwpdb.org/>) was used to convert them to mmCIF format. All jobs were run on the Stanford high-performance computing cluster, Sherlock 2.0, using Rosetta 3.10 (2020.42).

Model Validation. The modeling convergence, mean pairwise heavy-atom r.m.s.d., was calculated using the Rosetta command:

```
drafter_error_estimation -s $PDBs -mute core -rmsd_nosuper true
--per_residue_convergence true
```

The results can be found in *Dataset S1*. The stereochemical and map-to-model scores were calculated using the pipeline (https://github.com/DasLab/CASP15_RNA_EM), which includes using MolProbity (59), Phenix cross-correlation scores, CC_{volume} , CC_{mask} , and CC_{peaks} (60), Q-score (42), and TEMPy for Mutual Information (MI) and segment-based Manders' overlap coefficient (SMOC) scores (61). All calculations were carried out using the sharpened map and default parameters for each program. The mean per-residue convergence and Q-score of each 10 model ensemble were then calculated, saved as B-factors on a representative structure, and visualized using ChimeraX (62) using in-house scripts. The average scores can be found in *SI Appendix, Figs. S4, S10, and S13* and all scores can be found in *Dataset S2*. Models from the EternaFold secondary structure for all BtCoV-HKU5 SL5 conformations and from library-based DMS M2-seq from BtCoV-HKU5

conformation 4 were found by these metrics to not fit in the map sufficiently well and hence were not considered further (*SI Appendix, Fig. S10*). Finally, the effect of refinement using ERRASER2 on these validation metrics was plotted using in-house scripts. All scripts can be found in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D).

Model Analysis. The distance between UUYGU hexaloops was defined as the distance between centroids of the C1' atoms of the hexaloop. The inter-helical angle was defined as follows. A vector representing the SL5c:SL5-stem stack was defined by minimizing the distance between this vector and all heavy atoms pointing away from SL5c toward SL5-stem. Likewise, a vector was defined for the SL5a:SL5b stack pointing toward SL5b. The angle of the SL5c-to-SL5-stem vector relative to the SL5a-to-SL5b vector was defined as the inter-helical angle, with clockwise defined as positive. The direction of view was defined with the SL5c-to-SL5-stem vector on top of the SL5a-to-SL5b vector. Hence, a parallel configuration would result in a 0° inter-helical angle and antiparallel would result in a 180° inter-helical angle. The hinge angle was similarly defined but the vector was defined by the atoms in the apical residues SL5a stem-loop after the hinge, and a second vector as the remaining residues in the SL5a:SL5b stack. An angle of 0° would be a perfect coaxial stack; a positive angle indicates bends toward SL5-stem and negative away from SL5-stem. See *SI Appendix, Fig. S5* for a pictorial representation of these angles. The exact residues used to define the vectors can be found in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D). For figures, the pixel size of the SARS-CoV-2 SL5-6 domains map was increased from 1 Å/pixel to 1.1 Å/pixel to match other maps and the geometry of RNA A-form helices. Figures were prepared using ChimeraX (62), and scripts can be found in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D).

Data, Materials, and Software Availability. Raw movies and particle stacks are deposited to the Electron Microscopy Public Image Archive (EMPIAR), cryo-EM maps are deposited in the Electron Microscopy Data Bank (EMDB), and atomic models are deposited in the Protein Data Bank (PDB) with their accession numbers in *SI Appendix, Table S6* or in the accompanying GitHub repository (https://github.com/DasLab/Coronavirus_SL5_3D) (63). Reactivity traces are deposited in the RNA Mapping DataBase (RMDb) and sequencing data in the NIH Sequence Read Archive with the BioProject accession number: PRJNA1039878 (64). Repository Sites: SRA, RMDb, EMPIAR, EMDb, PDB (SRA: [SRR26810683](https://www.ncbi.nlm.nih.gov/sra/SRR26810683) (65), [SRR26810682](https://www.ncbi.nlm.nih.gov/sra/SRR26810682) (66), [SRR26810681](https://www.ncbi.nlm.nih.gov/sra/SRR26810681) (67), [SRR26810680](https://www.ncbi.nlm.nih.gov/sra/SRR26810680) (68), [SRR26827601](https://www.ncbi.nlm.nih.gov/sra/SRR26827601) (69) RMDb: [COVSL5_DMS_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (70), [COVSL5_DMS_0002](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (71), [COVSL5_NOM_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (72), and [COVSL5_NOM_0002](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (73), [SL5HKU_DMS_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (74), [SL5HKU_2A3_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (75), [SL5HKU_NOM_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (76), [SL5HKU_NOM_0002](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (77), [SL5MER_DMS_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (78), [SL5MER_2A3_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (79), [SL5MER_NOM_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (80), [SL5MER_NOM_0002](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (81), [SL5CV2_DMS_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (82), [SL5CV2_2A3_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (83), [SL5CV2_NOM_0001](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (84), [SL5CV2_NOM_0002](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1039878) (85) EMPIAR: [11827](https://www.ebi.ac.uk/EMPIAR/entry/11827) (86), [11813](https://www.ebi.ac.uk/EMPIAR/entry/11813) (87), [11834](https://www.ebi.ac.uk/EMPIAR/entry/11834) (88), [11814](https://www.ebi.ac.uk/EMPIAR/entry/11814) (89), [11838](https://www.ebi.ac.uk/EMPIAR/entry/11838) (90), [11815](https://www.ebi.ac.uk/EMPIAR/entry/11815) (91), [11837](https://www.ebi.ac.uk/EMPIAR/entry/11837) (92), [11836](https://www.ebi.ac.uk/EMPIAR/entry/11836) (93), [11835](https://www.ebi.ac.uk/EMPIAR/entry/11835) (94), [11848](https://www.ebi.ac.uk/EMPIAR/entry/11848) (95) = EMDb: [EMD-42818](https://www.ebi.ac.uk/EMDB/entry/EMD-42818) (96), [EMD-42821](https://www.ebi.ac.uk/EMDB/entry/EMD-42821) (97), [EMD-42820](https://www.ebi.ac.uk/EMDB/entry/EMD-42820) (98), [EMD-42819](https://www.ebi.ac.uk/EMDB/entry/EMD-42819) (99), [EMD-42816](https://www.ebi.ac.uk/EMDB/entry/EMD-42816) (100), [EMD-42809](https://www.ebi.ac.uk/EMDB/entry/EMD-42809) (101), [EMD-42810](https://www.ebi.ac.uk/EMDB/entry/EMD-42810) (102), [EMD-42811](https://www.ebi.ac.uk/EMDB/entry/EMD-42811) (103), [EMD-42801](https://www.ebi.ac.uk/EMDB/entry/EMD-42801) (104), [EMD-42805](https://www.ebi.ac.uk/EMDB/entry/EMD-42805) (105), [EMD-42802](https://www.ebi.ac.uk/EMDB/entry/EMD-42802) (106), [EMD-42808](https://www.ebi.ac.uk/EMDB/entry/EMD-42808) (107), [EMD-42803](https://www.ebi.ac.uk/EMDB/entry/EMD-42803) (108), [EMD-42813](https://www.ebi.ac.uk/EMDB/entry/EMD-42813) (109), [EMD-42814](https://www.ebi.ac.uk/EMDB/entry/EMD-42814) (110) PDB: [8UYS](https://www.rcsb.org/entry/8UYS) (111), [8UYU](https://www.rcsb.org/entry/8UYU) (112), [8UYK](https://www.rcsb.org/entry/8UYK) (113), [8UYL](https://www.rcsb.org/entry/8UYL) (114), [8UYM](https://www.rcsb.org/entry/8UYM) (115), [8UYE](https://www.rcsb.org/entry/8UYE) (116), [8UYG](https://www.rcsb.org/entry/8UYG) (117), [8UYJ](https://www.rcsb.org/entry/8UYJ) (118).

ACKNOWLEDGMENTS. We thank Paul Berg, Jeffrey S. Glenn, and members of the Glenn, Barna, and Puglisi labs (Stanford) for helpful discussions on the nature and structure of SL5. We thank Stanford University, Stanford Research Computing Center, SLAC Shared Scientific Data Facility, the Stanford-SLAC Cryo-EM Center, and the Stanford Biochemistry administrative staff for their computational resources and support that contributed to this research. This work was supported by Stanford BioX (Bowes Graduate Student Fellowship to R.C.K., Interdisciplinary Initiative Program to R.D.), the NIH (R35 GM122579 to R.D.), the NIH Antiviral Drug Discovery (AViDD) Centers for Pathogens of Pandemic Concern Program (U19 AI171421 to J.S. Glenn, W.C., and R.D.), Howard Hughes Medical Institute (HHMI) (to R.D.), the NSF (GRFP DGE-1656518 to L.X.), the NIH Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (U24 GM129541 to W.C.) for cryo-EM data collection, National Key R&D Program of China (2022YFC2303700 to K.Z. and S.L., and 2022YFA1302700 to K.Z.). This project was initiated during the COVID-19 pandemic with support from a Stanford University ChEM-H COVID-19

Drug and Vaccine Prototyping seed grant (P. Berg, J.S. Glenn, R.D., and W.C.). This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sub-licensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

- Zhou, Y. Qiu, X. Ge, The taxonomy, host range and pathogenicity of coronaviruses and other viruses in the Nidovirales order. *Animal Dis.* **1**, 1–28 (2021).
- D.-K. Kim *et al.*, A proteome-scale map of the SARS-CoV-2-human contactome. *Nat. Biotechnol.* **41**, 140–149 (2022).
- H. Yang, Z. Rao, Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat. Rev. Microbiol.* **19**, 685–700 (2021).
- M. L. Boby *et al.*, Open science discovery of potent noncovalent SARS-CoV-2 main protease inhibitors. *Science* **382**, eabo7201 (2023).
- I. Manfredonia *et al.*, Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.* **48**, 12436–12452 (2020).
- L. Sun *et al.*, In vivo structural characterization of the SARS-CoV-2 RNA genome identifies host proteins vulnerable to repurposed drugs. *Cell* **184**, 1865–1883.e20 (2021).
- N. C. Huston *et al.*, Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol. Cell* **81**, 584–598.e5 (2021).
- C. Iserman *et al.*, Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell* **80**, 1078–1091 (2020).
- R. Rangan *et al.*, De novo 3D models of SARS-CoV-2 RNA elements from consensus experimental secondary structures. *Nucleic Acids Res.* **49**, 3092–3108 (2021).
- T. C. T. Lan *et al.*, Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat. Commun.* **13**, 1–14 (2022).
- P. R. Bhatt *et al.*, Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science* **372**, 1306–1313 (2021).
- K. Zhang *et al.*, Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Nat. Struct. Mol. Biol.* **28**, 747–754 (2021).
- C. Roman, A. Lewicka, D. Koirala, N.-S. Li, J. A. Piccirilli, The SARS-CoV-2 programmed –1 ribosomal frameshifting element crystal structure solved to 2.09 Å using chaperone-assisted RNA crystallography. *ACS Chem. Biol.* **16**, 1469–1481 (2021).
- J. Vögele *et al.*, High-resolution structure of stem-loop 4 from the 5'-UTR of SARS-CoV-2 solved by solution state NMR. *Nucleic Acids Res.* **51**, 11318–11331 (2023).
- C. P. Jones, A. R. Ferré-D'Amaré, Crystal structure of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) frameshifting pseudoknot. *RNA* **28**, 239–249 (2022).
- V. Lulla *et al.*, Targeting the conserved stem loop 2 motif in the SARS-CoV-2 genome. *J. Virol.* **95**, e0066321 (2021).
- N. Vankadari, N. N. Jayasankar, W. J. Lopes, Structure of the SARS-CoV-2 Nsp1/5'-untranslated region complex and implications for potential therapeutic targets, a vaccine, and virulence. *J. Phys. Chem. Lett.* **11**, 9659–9668 (2020).
- C. W. Lee, L. Li, D. P. Giedroc, The solution structure of coronavirus stem-loop 2 (SL2) reveals a canonical CUYG tetraloop fold. *FEBS Lett.* **585**, 1049–1053 (2011).
- R. J. Hagey *et al.*, Programmable antivirals targeting critical conserved viral RNA secondary structures from influenza A virus and SARS-CoV-2. *Nat. Med.* **28**, 1944–1955 (2022).
- H. S. Haniff *et al.*, Targeting the SARS-CoV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RIBOTAC) degraders. *ACS Cent. Sci.* **6**, 1713–1721 (2020).
- M. Zafferani *et al.*, Amilorides inhibit SARS-CoV-2 replication in vitro by targeting RNA structures. *Sci. Adv.* **7**, eabl6096 (2021).
- S.-J. Park, Y.-G. Kim, H.-J. Park, Identification of RNA pseudoknot-binding ligand that inhibits the -1 ribosomal frameshifting of SARS-coronavirus by structure-based virtual screening. *J. Am. Chem. Soc.* **133**, 10094–10100 (2011).
- C. Becquart *et al.*, Exploring heterocycle-spermine conjugates as modulators of oncogenic microRNAs biogenesis. *ACS Omega* **3**, 16500–16508 (2018).
- S.-C. Chen *et al.*, Optimization of 5'UTR to evade SARS-CoV-2 Nonstructural protein 1-directed inhibition of protein synthesis in cells. *Appl. Microbiol. Biotechnol.* **107**, 2451–2468 (2023).
- R. Rangan *et al.*, RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: A first look. *RNA* **26**, 937–959 (2020).
- R. Madhugiri, M. Fricke, M. Marz, J. Ziebuhr, RNA structure analysis of alphacoronavirus terminal genome regions. *Virus Res.* **194**, 76–89 (2014).
- S. C. Chen, R. C. L. Olsthoorn, C. H. Yu, Structural phylogenetic analysis reveals lineage-specific RNA repetitive structural motifs in all coronaviruses and associated variations in SARS-CoV-2. *Virus Evol.* **7**, veab021 (2021).
- D. Yang, P. Liu, E. V. Wudeck, D. P. Giedroc, J. L. Leibowitz, 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 (2015).
- R. C. Kretsch *et al.*, RNA target highlights in CASP15: Evaluation of predicted models by structure providers. *Proteins* **91**, 1600–1615 (2023).
- D. Escors, A. Izeta, C. Capiscol, L. Enjuanes, Transmissible gastroenteritis coronavirus packaging signal is located at the 5' end of the virus genome. *J. Virol.* **77**, 7890–7902 (2003).
- K. Leppek *et al.*, Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat. Commun.* **13**, 1536 (2022).
- M. M. Hatmal *et al.*, Comprehensive structural and molecular comparison of spike proteins of SARS-CoV-2, SARS-CoV and MERS-CoV, and their interactions with ACE2. *Cells* **9**, 2638 (2020).
- M. E. Cueno, K. Imai, Structural comparison of the SARS CoV 2 spike protein relative to other human-infecting coronaviruses. *Front. Med.* **7**, 594439 (2021).
- K. Gottipati, S. C. McNamee, J. Tipo, M. A. White, K. H. Choi, Structural basis for cloverleaf RNA-initiated viral genome replication. *Nucleic Acids Res.* **51**, 8850–8863 (2023).
- S. Sharma, G. Varani, NMR structure of Dengue West Nile viruses stem-loop B: A key cis-acting element for flavivirus replication. *Biochem. Biophys. Res. Commun.* **531**, 522–527 (2020).
- E. Lee *et al.*, Structures of flavivirus RNA promoters suggest two binding modes with NS5 polymerase. *Nat. Commun.* **12**, 2530 (2021).
- S. Yan, Q. Zhu, J. Hohl, A. Dong, T. Schlick, Evolution of coronavirus frameshifting elements: Competing stem networks explain conservation and variability. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2221324120 (2023).
- K. Kappel *et al.*, Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **17**, 699–707 (2020).
- C. Y. Cheng, W. Kladwang, J. Yesselman, R. Das, RNA structure inference through chemical mapping after accidental or intentional mutations. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9876–9881 (2017).
- H. K. Wayment-Steele *et al.*, RNA secondary structure packages evaluated and improved by high-throughput experiments. *Nat. Methods* **19**, 1234–1242 (2022).
- F.-C. Chou, P. Sripathidevong, S. M. Dibrov, T. Hermann, R. Das, Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat. Methods* **10**, 74–76 (2013).
- G. Pintilie *et al.*, Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).
- S. L. Bonilla, M. E. Sherlock, A. MacFadden, J. S. Kieft, A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure. *Science* **374**, 955–960 (2021).
- N. J. Baird *et al.*, A discrete structure of an RNA folding intermediate revealed by cryo-electron microscopy. *J. Am. Chem. Soc.* **132**, 16352–16353 (2010).
- T. M. Nakamura, Y. H. Wang, A. J. Zaig, J. D. Griffith, T. R. Cech, Relative orientation of RNA helices in a group 1 ribozyme determined by helix extension electron microscopy. *EMBO J.* **14**, 4849–4859 (1995).
- J. S. Reuter, D. H. Mathews, RNAstructure: Software for RNA secondary structure prediction and analysis. *BMC Bioinform.* **11**, 129 (2010).
- W. Kladwang, C. C. VanLang, P. Cordero, R. Das, Understanding the errors of SHAPE-directed RNA structure modeling. *Biochemistry* **50**, 8049–8056 (2011).
- C. W. Leonard *et al.*, Principles for understanding the accuracy of SHAPE-directed RNA structure modeling. *Biochemistry* **52**, 588–595 (2013).
- M. Chen, M. F. Schmid, W. Chiu, Improving resolution and resolvability of single-particle cryoEM structures using Gaussian mixture models. *Nat. Methods* **21**, 37–40 (2023). [10.1038/s41592-023-02082-9](https://doi.org/10.1038/s41592-023-02082-9).
- J. Schwab, D. Kimanius, A. Burt, T. Dendooven, S. H. W. Scheres, DynaMight: Estimating molecular motions with improved reconstruction from cryo-EM images. *bioRxiv* [Preprint] (2023). <https://doi.org/10.1101/2023.10.18.562877> (Accessed 20 October 2023).
- W. Kamel *et al.*, Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. *Mol. Cell* **81**, 2851–2867.e7 (2021).
- A. Lescoute, E. Westhof, Topology of three-way junctions in folded RNAs. *RNA* **12**, 83–93 (2006).
- W. He, J. San Emeterio, M. T. Woodside, S. Kirmizialtin, L. Pollack, Atomistic structure of the SARS-CoV-2 pseudoknot in solution from SAXS-driven molecular dynamics. *Nucleic Acids Res.* **51**, 11332–11344 (2023).
- S. A. McKinney *et al.*, Single-molecule studies of DNA and RNA four-way junctions. *Biochem. Soc. Trans.* **32**, 41–45 (2004).
- S. L. Bonilla, J. S. Kieft, The promise of cryo-EM to explore RNA structural dynamics. *J. Mol. Biol.* **434**, 167802 (2022).
- R. Das, A. M. Watkins, RiboDraw: Semiautomated two-dimensional drawing of RNA tertiary structure diagrams. *NAR Genom. Bioinform.* **3**, lqab091 (2021).
- S. Tian, J. D. Yesselman, P. Cordero, R. Das, Primerize: Automated primer assembly for transcribing non-coding RNA domains. *Nucleic Acids Res.* **43**, W522–W526 (2015).
- A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
- C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
- D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
- T. Cragolini *et al.*, TEMPy2: A Python library with improved 3D electron microscopy density-fitting and validation workflows. *Acta Crystallogr. D Struct. Biol.* **77**, 41–47 (2021).
- E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
- R. C. Kretsch, L. Xu, G. Nye, Code accompanying Coronavirus SL5 3D Structure Determination. https://github.com/DasLab/Coronavirus_SL5_3D. Deposited 20 November 2023.
- R. Huang, R. C. Kretsch, I. N. Zheludev, R. Das, Coronavirus 5' proximal stem-loop 5 mutate-and-map. NCBI BioProject. <https://www.ncbi.nlm.nih.gov/bioproject/?term=1039878>. Deposited 13 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, M2Seq of SARS-CoV-2 SL5-6 modified (SRR26810683). SRA. [https://www.ncbi.nlm.nih.gov/sra/SRX22507798\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX22507798[accn]). Deposited 13 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, M2Seq of SARS-CoV-2 SL5-6 (SRR26810682). SRA. [https://www.ncbi.nlm.nih.gov/sra/SRX22507799\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX22507799[accn]). Deposited 13 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, M2Seq of SARS-CoV-2 truncated SL5 modified (SRR26810681). SRA. [https://www.ncbi.nlm.nih.gov/sra/SRX22507800\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX22507800[accn]). Deposited 13 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, M2Seq of SARS-CoV-2 truncated SL5 (SRR26810680). SRA. [https://www.ncbi.nlm.nih.gov/sra/SRX22507801\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX22507801[accn]). Deposited 13 November 2023.
- R. Huang, R. C. Kretsch, R. Das, SARS-CoV-2, HKU5, MERS, HCoV-229E, HCoV-63E SL5 for Cryo-EM (SRR26827601). SRA. [https://www.ncbi.nlm.nih.gov/sra/SRX22524118\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX22524118[accn]). Deposited 14 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, SARS-CoV-2 SL5-6 modified. COVSL5_DMS_0001. https://rmdb.stanford.edu/detail/COVSL5_DMS_0001. Deposited 13 November 2023.
- R. C. Kretsch, I. N. Zheludev, R. Das, SARS-CoV-2 SL5 modified. COVSL5_DMS_0002. https://rmdb.stanford.edu/detail/COVSL5_DMS_0002. Deposited 13 November 2023.

72. R. C. Kretsch, I. N. Zheludev, R. Das, SARS-CoV-2 SL5-6 not modified. COVSL5_NOM_0001. https://rmdb.stanford.edu/detail/COVSL5_NOM_0001. Deposited 13 November 2023.
73. R. C. Kretsch, I. N. Zheludev, R. Das, SARS-CoV-2 SL5 not modified. COVSL5_NOM_0002. https://rmdb.stanford.edu/detail/COVSL5_NOM_0002. Deposited 13 November 2023.
74. R. Huang, R. C. Kretsch, R. Das, BtCoV-HKU5 SL5 library DMS-modified. SL5HKU_DMS_0001. https://rmdb.stanford.edu/detail/SL5HKU_DMS_0001. Deposited 13 November 2023.
75. R. Huang, R. C. Kretsch, R. Das, BtCoV-HKU5 SL5 library 2A3-modified. SL5HKU_2A3_0001. https://rmdb.stanford.edu/detail/SL5HKU_2A3_0001. Deposited 13 November 2023.
76. R. Huang, R. C. Kretsch, R. Das, BtCoV-HKU5 SL5 library 2A3-unmodified. SL5HKU_NOM_0001. https://rmdb.stanford.edu/detail/SL5HKU_NOM_0001. Deposited 13 November 2023.
77. R. Huang, R. C. Kretsch, R. Das, BtCoV-HKU5 SL5 library DMS-unmodified. SL5HKU_NOM_0002. https://rmdb.stanford.edu/detail/SL5HKU_NOM_0002. Deposited 13 November 2023.
78. R. Huang, R. C. Kretsch, R. Das, MERS SL5 library DMS-modified. SL5MER_DMS_0001. https://rmdb.stanford.edu/detail/SL5MER_DMS_0001. Deposited 13 November 2023.
79. R. Huang, R. C. Kretsch, R. Das, MERS SL5 library 2A3-modified. SL5MER_2A3_0001. https://rmdb.stanford.edu/detail/SL5MER_2A3_0001. Deposited 13 November 2023.
80. R. Huang, R. C. Kretsch, R. Das, MERS SL5 library 2A3-unmodified. SL5MER_NOM_0001. https://rmdb.stanford.edu/detail/SL5MER_NOM_0001. Deposited 13 November 2023.
81. R. Huang, R. C. Kretsch, R. Das, MERS SL5 library DMS-unmodified. RNA Mapping Database. https://rmdb.stanford.edu/detail/SL5MER_NOM_0002. Deposited 13 November 2023.
82. R. Huang, R. C. Kretsch, R. Das, SARS-CoV-2 SL5 library DMS-modified. SL5CV2_DMS_0001. https://rmdb.stanford.edu/detail/SL5CV2_DMS_0001. Deposited 13 November 2023.
83. R. Huang, R. C. Kretsch, R. Das, SARS-CoV-2 SL5 library 2A3-modified. SL5CV2_2A3_0001. https://rmdb.stanford.edu/detail/SL5CV2_2A3_0001. Deposited 13 November 2023.
84. R. Huang, R. C. Kretsch, R. Das, SARS-CoV-2 SL5 library 2A3-unmodified. SL5CV2_NOM_0001. https://rmdb.stanford.edu/detail/SL5CV2_NOM_0001. Deposited 13 November 2023.
85. R. Huang, R. C. Kretsch, R. Das, SARS-CoV-2 SL5 library DMS-unmodified. SL5CV2_NOM_0002. https://rmdb.stanford.edu/detail/SL5CV2_NOM_0002. Deposited 13 November 2023.
86. R. C. Kretsch *et al.*, Cryo-EM of SARS-CoV-2 5' proximal stem-loop 5 (11827). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11827/>. Deposited 20 December 2023.
87. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of SARS-CoV-2 5' proximal stem-loops 5-6 (11813). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11813/>. Deposited 20 December 2023.
88. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of SARS-CoV-2 5' proximal stem-loops 5-6 with SL5b extended (11834). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11834/>. Deposited 20 December 2023.
89. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of SARS-CoV-2 5' proximal stem-loops 5-6 with SL5c extended (11814). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11814/>. Deposited 20 December 2023.
90. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of SARS-CoV-2 5' proximal stem-loops 5-6 with SL6 extended and SL5a, SL5b, and SL5c removed (11838). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11838/>. Deposited 20 December 2023.
91. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of SARS-CoV-1 5' proximal stem-loop 5 (11815). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11815/>. Deposited 20 December 2023.
92. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of MERS 5' proximal stem-loop 5 (11837). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11837/>. Deposited 20 December 2023.
93. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of BtCoV-HKU5 5' proximal stem-loop 5 (11836). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11836/>. Deposited 20 December 2023.
94. R. C. Kretsch, L. Xu, W. Chiu, R. Das, Cryo-EM of HCoV-229E 5' proximal stem-loop 5 (11835). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11835/>. Deposited 20 December 2023.
95. R. C. Kretsch, L. Xu, X. Zhou, W. Chiu, R. Das, Cryo-EM of HCoV-NL63 5' proximal stem-loop 5 (11848). EMPIAR. <https://www.ebi.ac.uk/empair/EMPIAR-11848/>. Deposited 14 November 2023.
96. R. C. Kretsch *et al.*, SARS-CoV-2 5' proximal stem-loop 5 map (EMD-42818). EMD. <https://www.ebi.ac.uk/emdb/EMD-42818>. Deposited 14 November 2023.
97. R. C. Kretsch *et al.*, SARS-CoV-2 5' proximal stem-loops 5-6 map (EMD-42821). EMD. <https://www.ebi.ac.uk/emdb/EMD-42821>. Deposited 14 November 2023.
98. R. C. Kretsch, L. Xu, W. Chiu, R. Das, SARS-CoV-2 5' proximal stem-loops 5-6, SL5b extended map (EMD-42820). EMD. <https://www.ebi.ac.uk/emdb/EMD-42820>. Deposited 14 November 2023.
99. R. C. Kretsch, L. Xu, W. Chiu, R. Das, SARS-CoV-2 5' proximal stem-loops 5-6, SL5c extended map (EMD-42819). EMD. <https://www.ebi.ac.uk/emdb/EMD-42819>. Deposited 14 November 2023.
100. R. C. Kretsch, L. Xu, W. Chiu, R. Das, SARS-CoV-1 5' proximal stem-loop 5 map (EMD-42816). EMD. <https://www.ebi.ac.uk/emdb/EMD-42816>. Deposited 14 November 2023.
101. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 1 map (EMD-42809). EMD. <https://www.ebi.ac.uk/emdb/EMD-42809>. Deposited 14 November 2023.
102. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 2 map (EMD-42810). EMD. <https://www.ebi.ac.uk/emdb/EMD-42810>. Deposited 14 November 2023.
103. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 3 map (EMD-42811). EMD. <https://www.ebi.ac.uk/emdb/EMD-42811>. Deposited 14 November 2023.
104. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 1 map (EMD-42801). EMD. <https://www.ebi.ac.uk/emdb/EMD-42801>. Deposited 14 November 2023.
105. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 2 map (EMD-42805). EMD. <https://www.ebi.ac.uk/emdb/EMD-42805>. Deposited 14 November 2023.
106. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 3 map (EMD-42802). EMD. <https://www.ebi.ac.uk/emdb/EMD-42802>. Deposited 14 November 2023.
107. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 4 map (EMD-42808). EMD. <https://www.ebi.ac.uk/emdb/EMD-42808>. Deposited 14 November 2023.
108. R. C. Kretsch, L. Xu, W. Chiu, R. Das, HCoV-229E 5' proximal stem-loop 5 map (EMD-42803). EMD. <https://www.ebi.ac.uk/emdb/EMD-42803>. Deposited 14 November 2023.
109. R. C. Kretsch, L. Xu, X. Zhou, W. Chiu, R. Das, HCoV-NL63 5' proximal stem-loop 5 conformation 1 map (EMD-42813). EMD. <https://www.ebi.ac.uk/emdb/EMD-42813>. Deposited 14 November 2023.
110. R. C. Kretsch, L. Xu, X. Zhou, W. Chiu, R. Das, HCoV-NL63 5' proximal stem-loop 5 conformation 2 map (EMD-42814). EMD. <https://www.ebi.ac.uk/emdb/EMD-42814>. Deposited 14 November 2023.
111. R. C. Kretsch *et al.*, SARS-CoV-2 5' proximal stem-loop 5 model (8UYJ). PDB. <https://www.rcsb.org/structure/8uyj>. Deposited 14 November 2023.
112. R. C. Kretsch, L. Xu, W. Chiu, R. Das, SARS-CoV-1 5' proximal stem-loop 5 model (8UYK). PDB. <https://www.rcsb.org/structure/8uyk>. Deposited 14 November 2023.
113. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 1 model (8UYL). PDB. <https://www.rcsb.org/structure/8uyl>. Deposited 14 November 2023.
114. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 2 model (8UYM). PDB. <https://www.rcsb.org/structure/8uym>. Deposited 14 November 2023.
115. R. C. Kretsch, L. Xu, W. Chiu, R. Das, MERS 5' proximal stem-loop 5 conformation 3 model (8UYE). PDB. <https://www.rcsb.org/structure/8uye>. Deposited 14 November 2023.
116. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 1 model (8UYG). PDB. <https://www.rcsb.org/structure/8uyg>. Deposited 14 November 2023.
117. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 2 model (8UYJ). PDB. <https://www.rcsb.org/structure/8uyj>. Deposited 14 November 2023.
118. R. C. Kretsch, L. Xu, W. Chiu, R. Das, BtCoV-HKU5 5' proximal stem-loop 5 conformation 4 model (8UYI). PDB. <https://www.rcsb.org/structure/8uyi>. Deposited 14 November 2023.