SUPPLEMENTARY MATERIAL

RNA DIMERIZATION PLAYS A ROLE IN RIBOSOMAL FRAMESHIFTING OF THE SARS CORONAVIRUS

Daniella Ishimaru[1], Ewan P. Plant[2,a], Amy C. Sims[4], Boyd L. Yount, Jr.[4], Braden M. Roth[1], Nadukkudy V. Eldho[1,b], Gabriela C. Pérez-Alvarado[3], David W. Armbruster[1], Ralph S. Baric[4], Jonathan D. Dinman[5], Deborah R. Taylor[2] and Mirko Hennig[1]

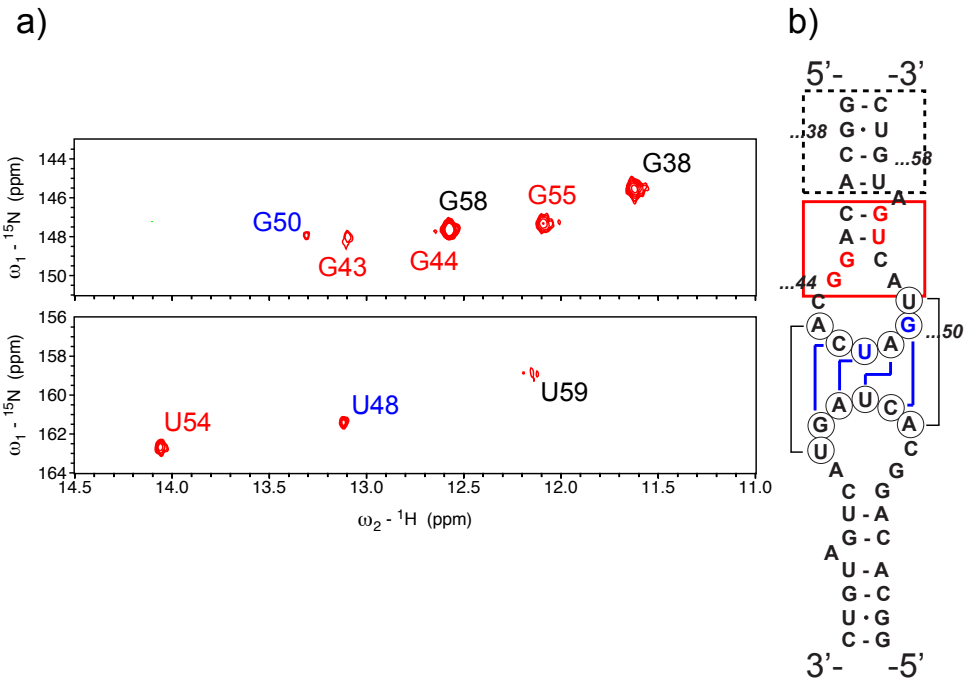
From the
[1]Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425,
[2]Laboratory of Hepatitis and Related Emerging Agents, Division of Emerging and Transfusion-Transmitted Diseases, Food and Drug Administration, Bethesda, Maryland 20892,
[3]Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, Illinois 62901,
[4]Departments of Epidemiology and [5]Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599 and the
[5]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742.
[a]Current address: Division of Viral Products, Office of Vaccine Research and Review, CBER, FDA, 8800 Rockville Pike, HFM445, Bethesda, MD 20892.
[b]Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742.

Address correspondence to:

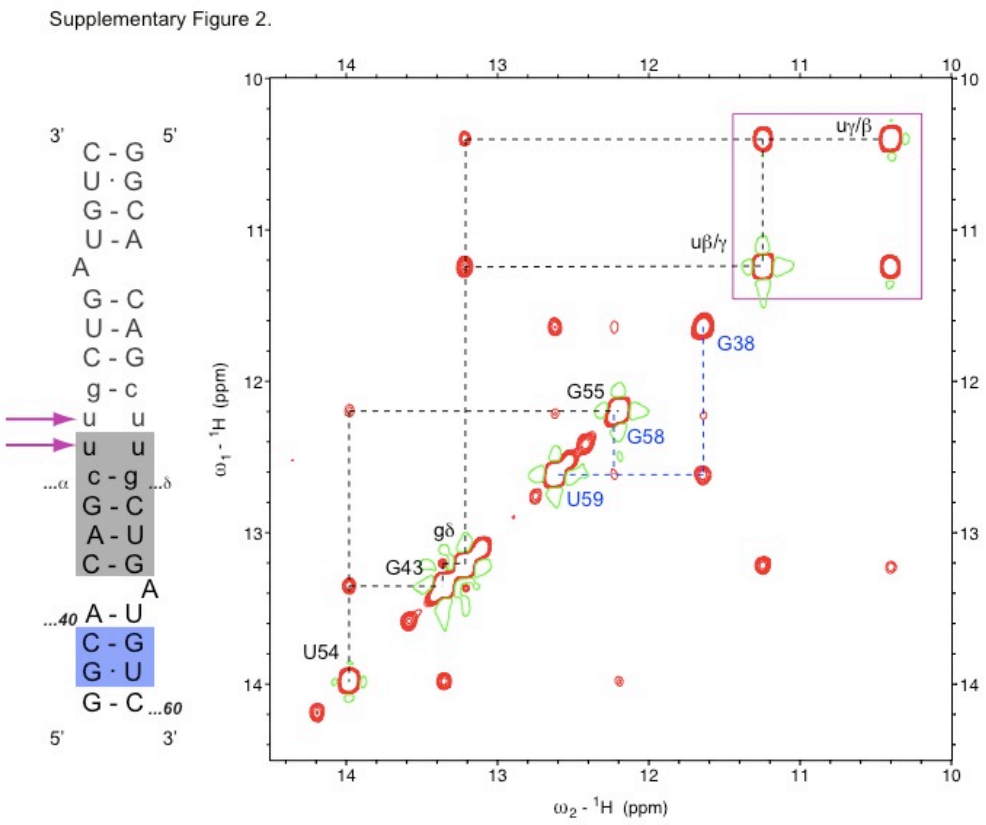
Mirko Hennig, Department of Biochemistry and Molecular Biology, Medical University of South Carolina, 70 President St, Charleston, South Carolina 29425. Email: hennig@musc.edu

Running Title: RNA dimerization of the SARS pseudoknot

Supplementary Figure 1.

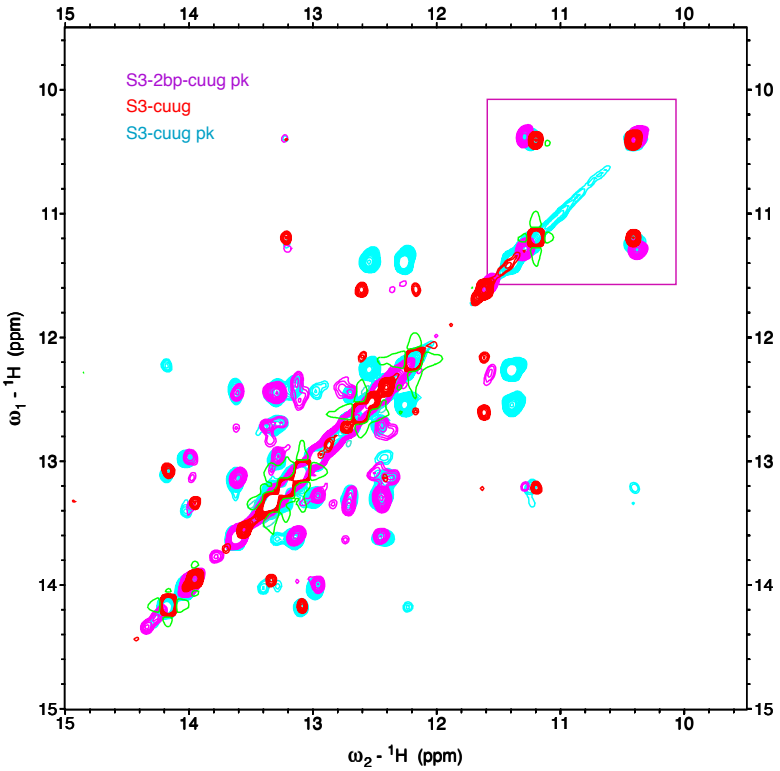


Supplementary Figure 1. Assignment of ^{1}H,^{15}N-HMQC NMR spectra of S3L2. (a) Imino regions of 2D ^{1}H,^{15}N-HMQC[58] experiments distinguish between guanosine (upper panel) and uridine (lower panel) iminos based on characteristic ^{15}N chemical shifts. Observable imino resonances are labeled with assignment information; while red labeled imino resonances are located in the upper portion of the stem, resonances labeled in black are found in the lower part of the S3 stem. Blue labels indicate imino resonances participating in kissing interactions. Note that G43 and U48 as well as G44 and G58 overlap in the ω_2-^{1}H imino proton dimension. Only the last two digits of the wild-type sequence numbering are used for clarity. Data were collected on a uniformly labeled ^{13}C,^{15}N-labeled SARS S3L2 containing 0.28 mM RNA in 500 µl volume of NMR buffer. Spectra were recorded at 298K on Bruker Avance III 600 MHz spectrometers equipped with a (H-C,N) triple resonance z-axis probe. **(b)** Schematic representation of a kissing-loop homodimer formed by S3L2. While the solid red box highlights cross-peaks in the upper part of the stem region, cross-peaks corresponding to the lower part of the stem are boxed using a black, dashed line. Solid blue lines connect imino protons involved in kissing interactions.



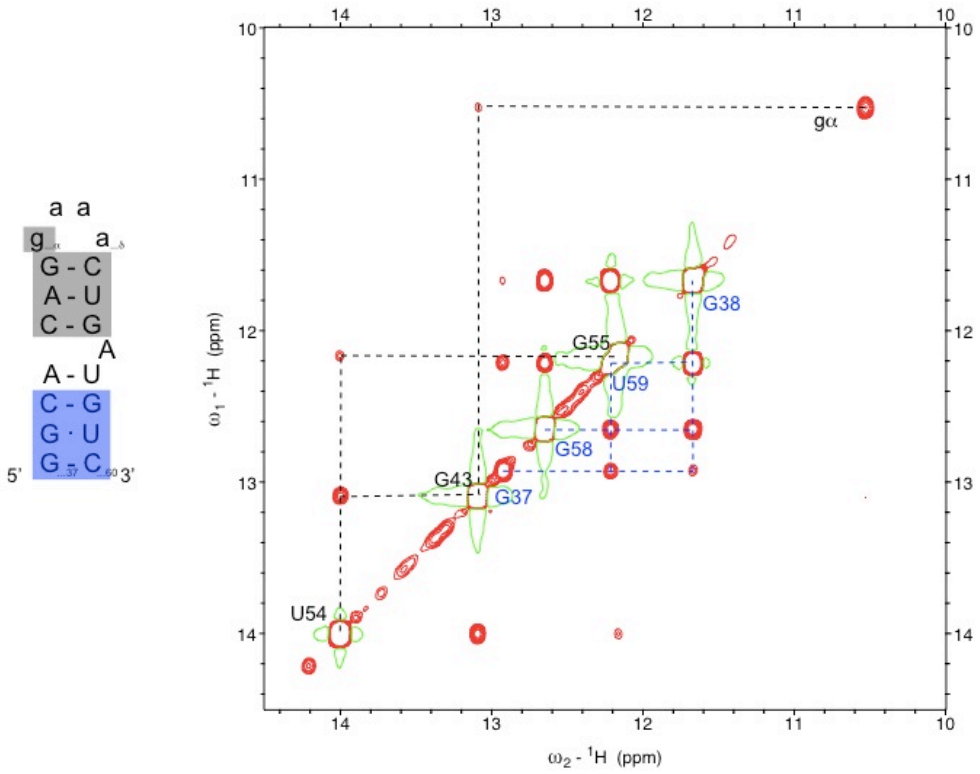
Supplementary Figure 2. NMR secondary structure of S3-cuug mutant. Imino region of 2D ^{1}H,^{1}H-NOESY experiment[51] with 200 ms mixing time collected on S3-cuug mutant. Spectrum was collected at 298K on Bruker Avance 800 MHz spectrometers equipped with cryogenically cooled triple resonance z-axis probe. Observable imino resonances are labeled with assignment information. Dashed black lines show the imino proton walk in upper part the S3-cuug stem. Dashed blue lines highlight the sequential cross peaks in the lower portion of S3-cuug correlating imino protons G38, U59, and G58. The solid magenta box highlights cross-peaks identifying the characteristic (tandem) Uβ/γ•Uγ/β wobble pair; arrows indicate the U•U position on the secondary structure. Data were collected on a SARS S3-cuug sample containing 0.6 mM RNA in 500 μl volume of NMR buffer. S3-cuug forms a stable, extended duplex as shown on the left.

Supplementary Figure 3.



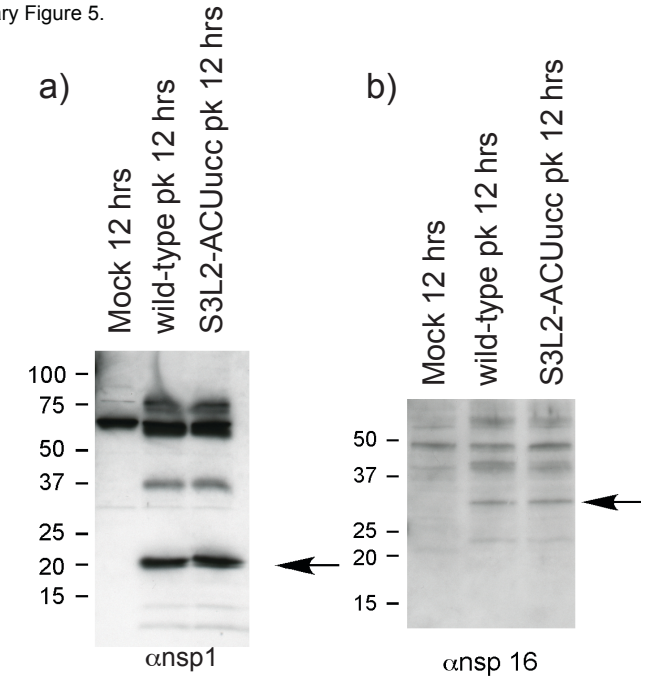
Supplementary Figure 3. Overlaid NMR spectra of S3-cuug, S3-2bp-cuug and S3L2-cuug pk mutants. Overlay of imino regions of 2D ^{1}H,^{1}H-NOESY experiments[51] with 200 ms mixing time collected on S3-cuug, S3-2bp-cuug and S3L2-cuug pk mutants. Spectra were collected at 298K on Bruker Avance 800 MHz spectrometers equipped with cryogenically cooled triple resonance z-axis probes. Data were collected on a SARS S3-cuug, S3-2bp-cuug and S3L2-cuug mutants samples containing 0.6-1.2 mM RNA in 500 µl volume of NMR buffer. All mutants form extended duplex structures. The solid magenta box highlights cross-peaks identifying the characteristic (tandem) U•U wobble pair.

Supplementary Figure 4.



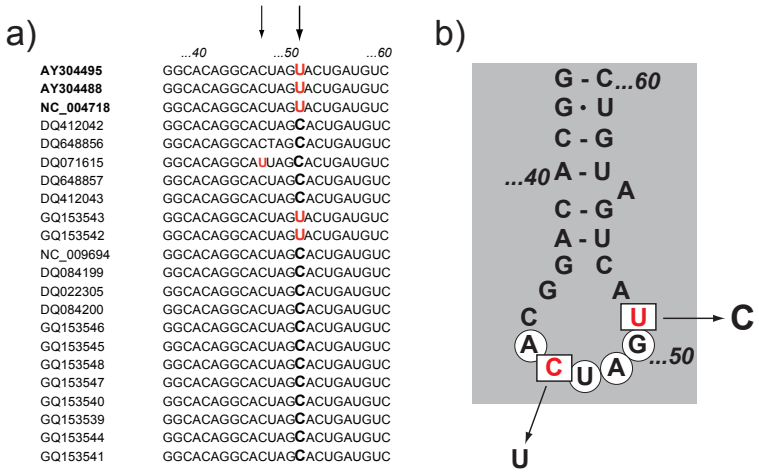
Supplementary Figure 4. NMR secondary structure of S3-gaaa mutant. Imino region of 2D ^{1}H,^{1}H-NOESY experiment[51] with 200 ms mixing time collected on S3-gaaa mutant. Spectrum was collected at 288K on Bruker Avance 800 MHz spectrometers equipped with cryogenically cooled triple resonance z-axis probe. Observable imino resonances are labeled with assignment information; only the last two digits of the wild-type sequence numbering are used for clarity. Data were collected on a SARS S3-gaaa sample containing 0.2 mM RNA in 500 μl volume of NMR buffer. S3-gaaa forms a stable hairpin as shown on the left. Dashed black lines show the imino proton walk in upper part the S3-gaaa stem including the observable loop gα imino proton. Dashed blue lines highlight the sequential cross peaks in the lower portion of S3-gaaa correlating imino protons G38, U59, and G58.

Supplementary Figure 5.

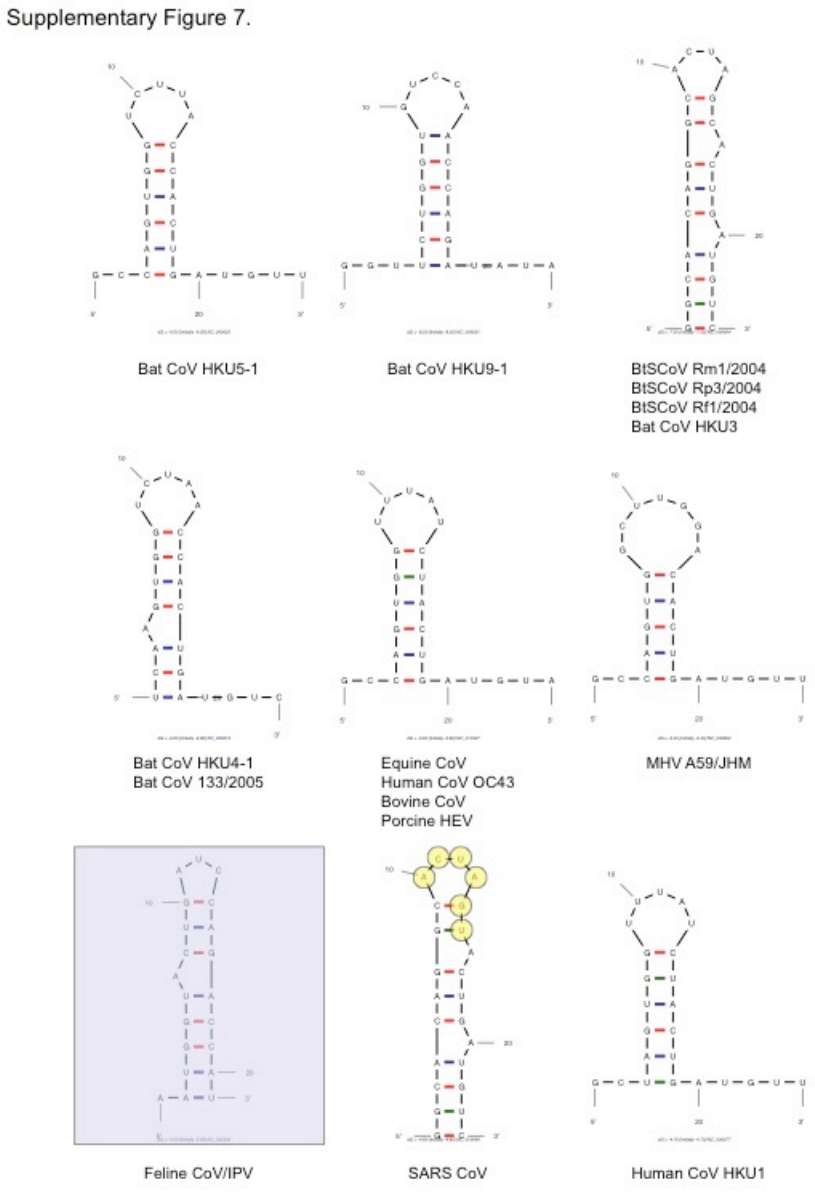


Supplementary Figure 5. ORF1a vs. ORF1b protein expression kinetics in S3L2-ACUucc frameshift mutant. Vero E6 cells were mock-infected or infected with wild type or S3L2-ACUucc SARS-CoV mutant viruses and harvested at 12 hours post infection. Proteins were separated on polyacrylamide gels and probed with rabbit sera directed against the replicase proteins nsp1 (**a**) or nsp16 (**b**) as indicated. Size markers are indicated to the left of each blot.

Supplementary Figure 6.



Supplementary Figure 6. Comparison of S3L2 stem-loop sequences of bat SL-CoV and SARS CoV. (a) Multiple sequence alignment of S3L2 stem-loops generated using ClustalW2. **(b)** Schematic representation of S3L2 with mutations disrupting the palindrome in SL-CoV indicated.



Supplementary Figure 7. Mfold-predicted secondary (sub)structures of CoV S3L2. Free energies ΔG of the proposed Stem 3-Loop 2 structures at 37°C in 1M NaCl were calculated with Mfold[28]. Resulting hairpin structures of CoV are ordered according to predicted thermodynamic stability with Bat CoV HKU5-1 being most stable (upper left corner) and Human CoV HKU1 being least stable (lower right corner). The only Group 1 CoV predicted to form a stable, structured S3 (Feline CoV/IPV) is boxed. The palindromic sequence of SARS CoV is highlighted using yellow circles.

Supplementary Table 1. Sequence Alignments of CoV frameshifting signals

Accession #	Name		S3 – Mfold ΔG[kcal/mol]	ClustalW2 alignment[a]
Group I				
NC_007025	Feline CoV	Ia	-5.6	TTTAAACGAGTGCGGGGTTCTAGTGCA---GCTCGACTAGAACCCTGT--- **AATGGTACTGATCCAGACCAT**GTTAGTAGAGCTTTT
NC_002306	Feline IPV	Ia	-5.6	TTTAAACGAGTGCGGGGTTCTAGTGCA---GCTCGACTAGAACCCTGC--- **AATGGTACTGATCCAGACCAT**GTTAGTAGAGCTTTT
NC_005831	Human CoV NL63	Ib	+0.4	TTTAAACGAGCAAGGGGTTCTAGTGCA---GCTCGACTAGAACCCTGC--- **AATGGCACGGACATCGATAAG**TGTGTTCGTGCTTTT
NC_009988	Bat CoV HKU2	Ib	-3.1	TTTAAACGGGCAAGGGGCTCTAGTGCA---GCTCGACTAGAGCCCTGT--- **AATGGTACTGAACCAGAACAT**TGTGTTCGTGCTTTT
NC_002645	Human CoV 229E	Ib	-0.2	TTTAAACGAGTCCGGGGCTCTAGTGCC---GCTCGACTAGAGCCCTGT--- **AATGGTACAGACATAGATTAC**TGTGTCCGTGCATTT
NC_010437	Bat coronavirus 1A	Ic	-2.4	TTTAAACGAGCAAGGGGCTCTAGTGCA---GCTCGACTAGAGCCCTGT--- **AATGGTACTGAACCGGAGCAT**GTTGTACGCGCCTTT
NC_010436	Bat coronavirus 1B	Ic	-3.1	TTTAAACGAGCAAGGGGCTCTAGTGCA---TCTCGACTAGAGCCCTGT--- **AATGGTACTGAACCAGAGCAT**GTTGTACGCGCCTTT
NC_010438	Bat CoV HKU8	Ic	-3.1	TTTAAACGAGTGCGGGGCTCTAGTGCA---GCTCGACTAGAGCCCTGT--- **AATGGTACTGAACCAGAACAC**GTAATCCGTGCCTTT
NC_009657	Bat CoV 512	Id	+0.5	TTTAAACGAGTGCGGGGCTCTAGTGCA---GCTCGACTAGAGCCCTTA--- **AATGGTTCGGACACACATCAT**GTGTTCCGTGCTTTT
NC_003436	Porcine EDV	Id	-0.2	TTTAAACGAGTACGGGGCTCTAGTGCA---GCTCGACTAGAGCCCTGT--- **AATGGTACTGATACACAACAT**GTGTATCGTGCTTTT
Group II				
NC_006577	Human CoV HKU1	IIa	-4.7	TTTAAACGGGTTCGGGGTACTAGTGTGAATGCCCGGCTAGTACCCTGT**GCTAGTGGTTTA** **TCTACTGATGTT**CAATTAAGGGCATTT
NC_001846	MHV strain A59	IIa	-6.4	TTTAAACGGGTTCGGGGTACAAGTGTAAATGCCCGTCTTGTACCCTGT**GCCAGTGGCTTG** **GACACTGATGTT**CAATTAAGGGCATTT
NC_006852	MHV strain JHM	IIa	-6.4	TTTAAACGGGTTCGGGGTACAAGTGTAAATGCCCGTCTTGTACCCTGT**GCCAGTGGCTTG** **GACACTGATGTT**CAATTAAGGGCATTT
NC_007732	Porcine HEV	IIa	-6.8	TTTAAACGGGTTCGGGGTACGAGTGTAGATGCCCGTCTCGTACCCTGT**GCCAGTGGTTTA** **TCTACTGATGTA**CAATTAAGGGCATTT
NC_003045	Bovine CoV	IIa	-6.8	TTTAAACGGGTTCGGGGTACGAGTGTAGATGCCCGTCTCGTACCCTGT**GCCAGTGGTTTA** **TCTACTGATGTA**CAATTAAGGGCATTT
NC_005147	Human CoV OC43	IIa	-6.8	TTTAAACGGGTTCGGGGTACGAGTGTAGATGCCCGTCTCGTACCCTGC**GCCAGTGGTTTA** **TCTACTGATGTA**CAATTAAGGGCATTT
NC_010327	Equine CoV	IIa	-6.8	TTTAAACGGGTTCGGGGTACGAGTGTAGATGCCCGTCTCGTACCCTGT**GCCAGTGGTTTA** **TCTACTGATGTA**CAATTAAGGGCATTT
NC_004718	SARS CoV	IIb	-4.8	TTTAAACGGGTTTGCGGTGTAAGTGCA--- GCCCGTCTTACACCGTGC**GGCACAGGC**ACTAGT**ACTGATGTC**GTCTACAGGGCTTTT
NC_013664	SARS CoV Rs672/2006	IIb	-4.8	TTTAAACGGGTTTGCGGTGTAAGTGCA--- GCCCGTCTTACACCGTGC**GGCACAGGC**ACTAGT**ACTGATGTC**GTCTACAGGGCTTTT
NC_009696	BtSCoV Rm1/2004	IIb	-7.5	TTTAAACGGGTTTGCGGTGTAAGTGCG--- GCCCGTCTTACACCGTGC**GGCACAGGCACTAGCACTGATGTC**GTTTACAGGGCTTTT
NC_009693	BtSCoV Rp3/2004	IIb	-7.5	TTTAAACGGGTTTGCGGTGTAAGTGCA--- GCCCGTCTTACACCGTGC**GGCACAGGCATTAGCACTGATGTC**GTCTACAGGGCTTTT
NC_009695	BtSCoV Rf1/2004	IIb	-7.5	TTTAAACGGGTTTGCGGTGTAAGTGCA--- GCCCGTCTTACACCGTGC**GGCACAGGCACTAGCACTGATGTC**GTTTACAGGGCTTTT
NC_009694	Bat CoV HKU3	IIb	-7.5	TTTAAACGGGTTTGCGGTGTAAGTGCG--- GCCCGTCTTACACCGTGC**GGCACAGGCACTAGCACTGATGTC**GTTTATAGGGCTTTT

NC_009019	Bat CoV HKU4-1	IIc	-6.8	TTTAAACGAGTCCGGGGTTCTAGTGTAAATGCCCGACTAGAACCCTGT**TCAAGTGGTCTA** **ACCACTGATGTC**GTTTATAGGGCATTC
NC_008315	Bat CoV 133/2005	IIc	-6.8	TTTAAACGAGTCCGGGGTTCTAGTGTAAATGCCCGACTAGAACCCTGT**TCAAGTGGTCTA** **ACCACTGATGTC**GTTTATAGGGCATTC
NC_009020	Bat CoV HKU5-1	IIc	-9.5	TTTAAACGAGTCCGGGGTTCTATTGTAAATGCCCGAATAGAACCCTGT**GCCAGTGGTCTT** **ACCACTGATGTT**GTCTTTAGGGCATTT
NC_009021	Bat CoV HKU9-1	IId	-8.2	TTTAAACGAGTCAGGGGTACTAGTGGAGTAGCCCGTCTAGTACCCCTA**GGTTCTGGTGTC** **CAACCAGATATA**GTATTAAGGGCTTTT
Group III				
NC_001451	Avian IBV	IIIa	-1.8	TTTAAACGGGTACGGGGTAGCAGTGAG--- GCTCGGCTGATACCCCTT**GCTAGTGGATGTGATCCTGATGTT**GTAAAGCGAGCCTTT

[a]Stem 3 – Loop2 regions examined by Mfold are shown in bold, the hexanucleotide palindrome observed in SARS CoV is highlighted in yellow.