Achieving a Golden Mean: Mechanisms by Which Coronaviruses Ensure Synthesis of the Correct Stoichiometric Ratios of Viral Proteins

Ewan P. Plant, † Rasa Rakauskaitė, ‡ Deborah R. Taylor, and Jonathan D. Dinman

Laboratory of Hepatitis and Related Emerging Agents, Division of Emerging and Transfusion-Transmitted Diseases, Office of Blood Research and Review, CBER, FDA, 8800 Rockville Pike, HFM310, Bethesda, Maryland 20892, and Department of Cell Biology and Molecular Genetics, Microbiology Building, Room 2135, University of Maryland, College Park, Maryland 20742

Received 24 November 2009/Accepted 4 February 2010

In retroviruses and the double-stranded RNA totiviruses, the efficiency of programmed −1 ribosomal frameshifting is critical for ensuring the proper ratios of upstream-encoded capsid proteins to downstream-encoded replicase enzymes. The genomic organizations of many other frameshifting viruses, including the coronaviruses, are very different, in that their upstream open reading frames encode nonstructural proteins, the frameshift-dependent downstream open reading frames encode enzymes involved in transcription and replication, and their structural proteins are encoded by subgenomic mRNAs. The biological significance of frameshifting efficiency and how the relative ratios of proteins encoded by the upstream and downstream open reading frames affect virus propagation has not been explored before. Here, three different strategies were employed to test the hypothesis that the −1 PRF signals of coronaviruses have evolved to produce the correct ratios of upstream- to downstream-encoded proteins. Specifically, infectious clones of the severe acute respiratory syndrome (SARS)-associated coronavirus harboring mutations that lower frameshift efficiency decreased infectivity by >4 orders of magnitude. Second, a series of frameshift-promoting mRNA pseudoknot mutants was employed to demonstrate that the frameshifting signals of the SARS-associated coronavirus and mouse hepatitis virus have evolved to promote optimal frameshift efficiencies. Finally, we show that a previously described frameshift attenuator element does not actually affect frameshifting per se but rather serves to limit the fraction of ribosomes available for frameshifting. The findings of these analyses all support a “golden mean” model in which viruses use both programmed ribosomal frameshifting and translational attenuation to control the relative ratios of their encoded proteins.

Viruses utilize programmed ribosomal frameshifting (PRF) to posttranscriptionally regulate the expression of multiple genes encoded on monocistronic viral mRNAs. In many RNA viruses that utilize programmed ribosomal frameshifting (e.g., most retroviruses, totiviruses, and Ty elements), the mRNAs transcribed from these viral templates contain two overlapping open reading frames (ORFs). In these viruses, the ORF encoding the major viral nucleocapsid proteins (e.g., Gag) is located at the 5′ end of the mRNA, whereas ORFs encoding proteins with enzymatic functions (typically Pro and Pol) are located 3′ of, and out of frame with, the Gag ORF. The enzymatic proteins are translated only as a result of PRF events that occur at frequencies of 1 to 40% depending on the specific virus and assay system employed (reviewed in reference 6). Thus, the majority of translational events result in the production of structural nucleocapsid proteins, while the intervention of frameshifting results in a decreased yield of enzymatic products (23). The importance of maintaining precise ratios of structural to enzymatic proteins on viral propagation has been demonstrated using two endogenous viruses of the yeast Saccharomyces cerevisiae and with two retroviruses (reviewed in reference 18). Small alterations in programmed frameshifting efficiencies promote the rapid loss of the yeast double-stranded RNA (dsRNA) L-A killer virus (13, 14, 17, 19, 38, 39, 40, 44, 49). Similarly, increasing or decreasing the efficiency of the +1 ribosomal frameshift in the TyI retrotransposable element of yeast results in reduced retrotransposition frequencies (2, 17, 20, 27, 28, 33, 39). In L-A, Gag-pol dimerization nucleates the formation of the virus particles (10–12, 22). Increasing the amount of Gag-pol protein synthesized may cause too many particles to initiate nonproductively, while producing too little may prevent efficient dimerization (19). The proteolytic processing of the TyA-TyB (Gag-pol equivalent) polyprotein of TyI is more akin to the situation observed in retroviruses. In TyI, increasing the amount of Gag-pol protein synthesized inhibited the proteolytic processing of the polyprotein (33). As a consequence, the formation of the mature forms of RNase H, integrase, and reverse transcriptase is blocked (33). Similarly, changing the ratio of Gag to Gag-pol proteins in retroviruses like HIV or Moloney murine leukemia virus interferes with virus particle formation (4, 24, 29, 32, 42, 53). In these cases, the overexpression of the Gag-pol protein results in the inefficient processing of the polyprotein and the inhibition of virus production. In sum, viral PRF efficiencies

† Present address: Division of Viral Products, Office of Vaccine Research and Review, CBER, FDA, 8800 Rockville Pike, HFM445, Bethesda, MD 20892.
‡ Published ahead of print on 17 February 2010.

* Corresponding author. Mailing address: Department of Cell Biology and Molecular Genetics, Microbiology Bldg, Rm 2135, University of Maryland, College Park, Maryland 20742. Phone: (301) 405-0918. Fax: (301) 314-9489. E-mail: dinman@umd.edu.
have been fine-tuned to deliver the precise ratios of proteins required for efficient viral particle assembly; too much or too little frameshifting alters this ratio, with detrimental consequences. Based on these studies, it has been proposed that \(-1\) PRF is a viable target for the prevention of viral propagation (reviewed in 18).

Coronaviruses are positive-strand RNA viruses with large genomes (~30,000 nucleotides [nt]) that also utilize genomes (\(/H11011\)) (reviewed in 18). PRF is a viable target for the prevention of viral propagation sequences. Based on these studies, it has been proposed that \(-1\) PRF is a viable target for the prevention of viral propagation (reviewed in 18).

Coronaviruses are positive-strand RNA viruses with large genomes (~30,000 nucleotides [nt]) that also utilize \(-1\) PRF. They can cause enteric and respiratory tract infections with varying severity. For example, some genotypes affecting humans (HCoV-229E and HCoV-OC43) cause cold-like symptoms, while the coronavirus associated with severe acute respiratory disease (SARS-CoV) is associated with a high mortality rate. Similarly, the coronaviruses that affect other mammals have assorted phenotypes: the mouse hepatitis virus (MHV) enterotropic strains replicate initially in the intestinal epithelium and tend not to disseminate, whereas the neurotrophic MHV strains initially replicate in the respiratory tract and then disseminate to the liver, brain, and lymph nodes. The latter strains are used in models for acute and chronic central nervous system infection (54). While the SARS-CoV and MHV viruses have different pathologies, overall they are phylogenetically more similar to each other than SARS-CoV is to HCoV-229E (21). The genomic organization of coronaviruses is different from that of retroviruses and totiviruses: the structural proteins are encoded by subgenomic mRNAs, while the genes regulated by \(-1\) PRF are involved in replicase/transcriptase function (56, 59). The genomic organization of SARS-CoV is shown in Fig. 1. The ORF1a-encoded polyprotein (pp1a) synthesizes nonstructural proteins. The \(-1\) PRF signal is located at the 3’ end of ORF1a and redirects a fraction of the translating ribosomes into the ORF1b reading frame to synthesize the larger pp1ab polyprotein. The enzymatic functions required for viral replication are derived from pp1ab (1, 5, 55). Although frameshifting is an essential feature of the viral life cycle per se because it is required for the production of most of the replicase proteins, the consequences of changing \(-1\) PRF efficiencies on the replication of this class of viruses have never been tested.

The cis-acting signals that promote frameshifting consist of a heptameric slippery site and an strong mRNA structure separated by a short spacer. In general, the slippery site can be defined as N NNW WWH, where N is any three identical bases, W is AAA or UUU, and H is A, C, or U (the frame of the initiator AUG is indicated by the spacing) (8, 16). It appears that there is a preference within virus groups for certain slippery sites, and these preferences likely reflect the differences in the host ribosomes (3, 45). The second element is usually an mRNA pseudoknot that directs elongating ribosomes to pause with their A and P sites positioned over the slippery site (34, 51). The initial demonstration that a pseudoknot was required for efficient \(-1\) PRF was for the avian infectious bronchitis coronavirus (IBV) (9). Subsequently, numerous pseudoknots have been described that facilitate frameshifting (reviewed in reference 25). Until recently, all of the frameshifting pseudoknots described contained two stems. However, structural analyses revealed that the SARS-CoV frameshift-stimulating pseudoknot contains three stems (47, 52). In addition, another cis-acting element affecting \(-1\) PRF located immediately upstream of the SARS-CoV \(-1\) PRF signal was suggested to attenuate the frameshifting efficiency of both the SARS-CoV and infectious bronchitis virus (IBV) signals (52). The availability of sequences from several new coronaviruses now allows more in-depth comparisons of regulatory sequences.

The current study begins by examining the question of the importance of synthesizing the correct ratios of viral proteins for SARS-CoV propagation and then addresses mechanisms through which these ratios may be controlled. Initially, a series of slippery-site mutants was introduced into an infectious clone to test the hypothesis that correct levels of \(-1\) PRF are critical for the propagation of this virus. The viable mutant viruses produced less genomic RNA than subgenomic RNA. Furthermore, the infection of cells with equivalent amounts of wild-type and mutant genomic RNAs revealed that the mutants were significantly less infectious than the wild type, thus demonstrating an important role for \(-1\) PRF in the viral life cycle. The hypothesis that the frameshift-stimulating mRNA pseudoknots have evolved in coronaviruses to promote frameshifting at specific levels so as to deliver the proper ratios of ORF1a and ORF1ab products was tested using a series of mutations that morphed the MHV \(-1\) PRF signal into that from SARS-CoV. The results of this analysis revealed features of the coronavirus pseudoknots that are important for stimulating optimal levels of frameshifting. Lastly, the issue of an additional regulatory element, the so-called attenuator sequence...
(52), was examined. Phylogenetic analyses reveal that while there is little conservation of the sequence upstream of the various coronavirus 1 PRF signals, computational analyses show that they all are predicted to fold into strong secondary structures. Although prior findings suggested that the attenuator element reduced 1 PRF by ~40%, the experimental design employed in that study did not preclude the hypothesis that strong secondary mRNA structures simply cause ribosomes to dissociate from the mRNA prior to encountering the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting that the function of the attenuator is to further help fine-tune the ratios of ORF1a- and ORF1b-tuned ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a- and ORF1b-tuned ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting the frameshift signal, i.e., translational attenuation. Experiments presented in the current study support this hypothesis, suggesting that the function of the attenuator is to further help fine-tune the ratios of ORF1a and ORF1b viral products by limiting

### MATERIALS AND METHODS

**Sequence analysis.** The GenBank (http://www.ncbi.nlm.nih.gov/GenBank/) accession numbers for the sequences discussed in this paper are SARS-CoV (NC_004718), Bt-CoV Rp3 (NC_009693), Bt-CoV HKU3 (NC_009694), Bt-CoV Rt1 (NC_009695), Bt-CoV HKU4 (NC_009696), Bt-CoV HKU-1 (NC_009691), Bt-CoV HKU5 (NC_009692), Bt-CoV HKU9 (NC_009690), Bt-CoV 133/2005 (NC_008315), BV (NC_001451), TCV (NC_001080), MvCoV (NC_011550), BtcCoV (NC_001548), ThCoV (NC_011549), SW1-CoV (NC_010686), MHV-As9 (NC_001846), MHV-HM (NC_006852), HCoV OC43 (NC_005147), BCoV (NC_003045), EqCoV (NC_010327), porcine epidemic diarrhea virus (NC_003436), Bt-CoV 512/2005 (NC_009657), Bt-CoV HKU8 (NC_010438), Bt-CoV 1a (NC_010437), and Bt-CoV 1b (NC_010436). Sequences were aligned using ClustalW2 (35), and clades were constructed on the EMBL website (http://www.ebi.ac.uk/). Pairwise alignments were performed using the default Clustal settings in the Lasergene software (DNASTAR Inc., Madison, WI). RNA sequences were folded using mfold on the web server at Rensselaer Polytechnic Institute (36, 60).

**Strains and genetic methods.** Escherichia coli strain DH5α was used to amplify plasmids, and high-efficiency transformations were performed using the method of Iino et al. (30). Vero E6 cells were cultured at 37°C with 5% CO2 in Dulbecco's modified eagle medium (Invitrogen, Carlsbad, CA) and the primers listed in Table 1. The mutations were confirmed by DNA sequencing.

### Table 1. Oligonucleotide primers used in this study

<table>
<thead>
<tr>
<th>Mutagenesis</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHVII</td>
<td>5'-AAGGATCCCTGTTCCTGTAGTGACC-3'</td>
</tr>
<tr>
<td>MHVII</td>
<td>5'-CCGGATCTTTTTAAAAGGGGTTCGG-3'</td>
</tr>
<tr>
<td>MHVIII</td>
<td>5'-CCGAGCCTAACCCCTTAATTGAAACTC-3'</td>
</tr>
<tr>
<td>MHVstem2</td>
<td>5'-AAGGTTTCGCCCTTTGTACC-3'</td>
</tr>
<tr>
<td>MHVloop3f</td>
<td>5'-TGATGTCTTAAAGGGCATTTTGAC-3'</td>
</tr>
<tr>
<td>MHVloop3r</td>
<td>5'-CTTAAAGACATGTTGCAACAGC-3'</td>
</tr>
<tr>
<td>MHVloop3bulge</td>
<td>5'-TGATGTCTACAGGGCATTTGACTC-3'</td>
</tr>
<tr>
<td>MHVloop3bulger</td>
<td>5'-GCCCTGTAGACATGTTGAAATGGTC-3'</td>
</tr>
<tr>
<td>MHVcomplex</td>
<td>5'-TGGCACATAGTCACATGGTGTAGTGCCAC-3'</td>
</tr>
<tr>
<td>MHVcomplex2</td>
<td>5'-CTTCTAGACACATTGCTAATGTCGACC-3'</td>
</tr>
</tbody>
</table>

**Attenuator analysis**

| T7KozakSARS          | 5'-TAATACGACTCACTATAGGAGGCCACCATTTATGATC-3' |
| RevSARS              | 5'-CCCTCATCTTAATGACGTAGAGC-3'               |
| T7SARS23kd           | 5'-TAATACGACTCACTATAGGAGGCCACCATTTTTTGAGAATAAC-3' |
| RevSAR23kd           | 5'-CCACCGCTAAAAGAATTGAAG-3'                 |
| T7Luc                | 5'-TATATACGACTCACTATAGGAGGCCACCATTTTTTGAGAATAAC-3' |
| revLuc               | 5'-GCCCTGTAGACATGTTGAAATGGTC-3'             |
| T7lorSHAPE           | 5'-TAATACGACTCACTATAGGAGGCCACCATTTTTTGAGAATAAC-3' |
| revSHAPE             | 5'-GCCCTGTAGACATGTTGAAATGGTC-3'             |
| FlucFor              | 5'-CCAAATGCTTGTGGAAGGTGC-3'                 |
| SHAPEoligo           | 5'-GCCCTGTAGACATGTTGAAATGGTC-3'             |
| TauMan analysis      | 5'-TGGCTTTTTTGGAAATGTCCTC-3'                |
| SARS13496            | 5'-AAAAAATTGTCCCTCCTCACTTTCCTC-3'           |
| SARS13564            | 5'-AAAAAATTGTCCCTCCTCACTTTCCTC-3'           |
| SARS30-50            | 5'-AAAAAATTGTCCCTCCTCACTTTCCTC-3'           |
| SARS28539            | 5'-AAAAAATTGTCCCTCCTCACTTTCCTC-3'           |

[M. S. Plant et al.](#)
SARS-CoV reverse genetics. Briefly, a full-length cDNA clone of the SARS-CoV genome was constructed from six subclones (called SARS clones A through F), and SP6 RNA polymerase and a GTP cap analog were used to generate full-length infectious transcripts (58). The translation of these into mammalian cell lines results in a productive, lytic viral infection. RNA was prepared by in vitro transcription, and the transcripts were transfected into Vero cells. The virus was allowed to grow for 5 days at 37°C. Viral supernatants were produced using the Applied Biosystems high-capacity cDNA reverse transcription kit (Foster City, CA) according to the manufacturer’s instructions. Primers complementary to nucleotides 13496 to 13516 and 13564 to 13542 (using the numbering of the Urbani SARS-CoV strain; GenBank numberAY277871) were used to detect genomic transcripts, and primers complementary to nucleotides 30 to 50 and 28539 to 28521 were used to detect subgenomic RNA. Although the primers for detecting the sgRNA can anneal to the gRNA, only the smallest sgRNA has the ability to translate given the mRNA cap structure (59). Primers complementary to nucleotides 13496 to 13516 and 13564 to 13542 (using the numbering of the Urbani SARS-CoV strain; GenBank numberAY277871) were used to detect genomic transcripts, and primers complementary to nucleotides 30 to 50 and 28539 to 28521 were used to detect subgenomic RNA. Although the primers for detecting the sgRNA can anneal to the gRNA, only the smallest sgRNA has the ability to translate given the mRNA cap structure (59).

Quantification of viral titer. The abundance of viral genomic and subgenomic RNAs (gRNA and sgRNA, respectively) was determined by quantitative real-time PCR using SYBR green chemistry. RNA was extracted from infected cells using TRIzol (Invitrogen, Carlsbad CA) from which cDNA was produced using the Applied Biosystems high-capacity cDNA reverse transcription kit (Foster City, CA) according to the manufacturer’s instructions. Primers complementary to nucleotides 13496 to 13516 and 13564 to 13542 (using the numbering of the Urbani SARS-CoV strain; GenBank numberAY277871) were used to detect genomic transcripts, and primers complementary to nucleotides 30 to 50 and 28539 to 28521 were used to detect subgenomic RNA. Although the primers for detecting the sgRNA can anneal to the gRNA, only the smallest sgRNA has the ability to translate given the mRNA cap structure (59). The forward primers included the T7 transcription promoter (Table 1). Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'.

Translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'. Small amplicons were generated so that differences in the proteins from translation assays. PCR primers were designed to amplify the frameshift cassette. PCR products were made using the Fermentas 2/1 H11003 and the primers T7 for SHAPE, 5'-TAATACGACTCACTATAGGGAAGA TGCACCTGATGAAATGG-3'.

RESULTS

Frameshift efficiency plays a critical role in SARS-CoV propagation. The composition of a slippery site can strongly influence its ability to promote −1 PRF (9, 19, 45). A series of slippery-site mutants was constructed in a dual luciferase-based −1 PRF reporter plasmid (pJD502), and the effects on −1 PRF efficiency were assessed as previously described (47). The plasmids are based on those described by Grentzmann et al. (26) and support transient expression. The activity of the second luciferase is dependent on −1 PRF and is normalized to the activity of the first. As shown in Table 2, the different slippery sites promoted a broad range of −1 PRF efficiencies from 14.4 to 0.15%. A reverse genetics system (58) was then employed to assess the effects of three of the mutant slippery sites (marked by an asterisk in Table 2) on SARS-CoV propagation. The three mutant viruses (and wild-type control) were constructed, recovered, and tested for infectivity using a tissue culture infectious dose assay. Initial visual inspection revealed that plaque numbers and diameters increased as −1 PRF efficiencies approached wild-type levels. Consistent with the hypothesis that −1 PRF efficiency is critical for coronavirus prop-
agitation, the U UUG AAC mutant, which does not alter the primary protein coding sequence of the slippery site but which almost completely abolished frameshifting, was not able to support virus propagation. RNAs from the wild-type and remaining two mutant viruses were recovered and quantified using real-time PCR. Interestingly, the abundance of genomic RNA in infected cells from the mutant viruses was reduced by approximately three orders of magnitude relative to wild-type RNA in infected cells from the mutant viruses were recovered and quantified almost completely abrogated frameshifting, was not able to propagation, the U UUG AAC mutant, which does not alter the RNA transcribed from a SARS replicon (1). The number of copies per ml of viral stock is shown with standard deviations.

The predicted structural differences between the SARS-CoV and MHV frameshift-promoting pseudoknots provided natural starting points to probe for how differences in structure affect function. To this end, a series of four mutants was constructed that sequentially changed the MHV −1 PRF-promoting pseudoknot into the SARS-CoV structure. Figure 3A shows the entire series of six constructs, from wild-type MHV to SARS-CoV, where the predicted secondary structures were computationally determined using pknots (48). The circles denote bases sequentially mutagenized to morph the MHV pseudoknot into the final SARS-CoV structure. To gain a better sense of their solution structures, SHAPE analysis of T7 RNA polymerase primer extension reactions was employed using NMIA (57). Representative autoradiograms of the SHAPE reactions for each construct are shown in Fig. 3B, and the results are color mapped onto the secondary structures in Fig. 3A, where black denotes fully protected bases, green means partial protection from modification, and red shows that the sugars were fully available for modification by NMIA. Most aspects of the predicted structures held true, with fully protected bases (black) mapping inside of the stems and partially protected bases (green) mapping to stem junctions and in loops. All of the stem 1 structures were highly stable. However, this analysis did reveal some important differences from the predicted structures. For example, although the MHV stem 2 is predicted to be 9 bp in length, the high A-U and G-U content at its distal end rendered it quite unstable, decreasing its actual length to a core of three to four G-C base pairs and consequentially increasing the length of loop 1 from the predicted 2 nt to approximately 7 nt. The SARS-CoV stem 2 was much more stable (5 to 6 bp plus the protected bulged A), and its actual loop 1 was shorter than that of MHV (4 nt). In contrast, while base pairing in the MHV stem 3 was very stable, progressive mutations toward the SARS-CoV sequence led base pairing in this region to become less stable. Further, while the repositioning of the bulged A in stem 3 helped to stabilize stem 2, this further destabilized stem 3. Similar patterns were observed in the loop 2 and loop 3 regions, although we note that loop changes simultaneous with moving the bulged A may have influenced stem 3 stability. For example, while the larger loop 2 of SARS-CoV was significantly more solvent accessible than its shorter counterpart in MHV, the longer loop 3 of MHV was more solvent accessible than the shorter loop 3 of SARS-CoV. Interestingly, increasing the stability of loop 3 by reducing its size progressively destabilized loop 2 (compare the series Δ2, Δ2Δ3, Δ2bΔ3 in Fig. 1A). In summary, while the MHV and SARS-CoV −1 PRF-promoting pseudoknots both contain stable stem 1 structures, they each appear to have exchanged solvent-accessible (stem 2)/loop 3 for MHV and stem 3/loop 2 for SARS-CoV) and solvent-inaccessible modules (stem 3/loop 2 for MHV and stem 2/loop 3 for SARS-CoV). In contrast, the intermediate constructs appear to progress to less-solvent-accessible, and thus more compacted structures, culminating with Δ2 and Δ2Δ3, from which point they sequentially become more solvent accessible and, thus, less compacted.
The effects of these constructs on −1 PRF efficiency were assayed as previously described (47). The wild-type MHV and SARS-CoV sequences promoted roughly equivalent levels of −1 PRF (SARS-CoV, ~18.4%; MHV, ~24.6%) (Fig. 3A). Beginning with the MHV sequence, shortening loop 1 by one codon (Δ2) increased the stability of stem 2 and increased frameshifting efficiency to ~66%. Reducing the length of loop 3 further compacted the structure (Δ2Δ3), increasing frameshifting to ~90.8%. The insertion of a bulge into stem 2 (Δ2bΔ3) rendered loop 3 more solvent accessible (less compacted) and lowered the frameshifting frequency to ~51.7%. A final mutant that repositioned the bulge in stem 3 made the loop more like that of the SARS-CoV pseudoknot (Δ2bΔ3b) lowered frameshifting efficiency back to a level closer to that of
the wild-type SARS-CoV basal level (∼26.8%). These results demonstrate a relationship between pseudoknot stability/compactness and −1 PRF efficiency and support the hypothesis that the pseudoknots of SARS-CoV and MHV evolved to promote the correct frequencies of −1 PRF required for the optimum synthesis of ORF1a- and ORF1b-encoded proteins.

The SARS-CoV attenuator sequence impedes ribosome processivity rather than inhibiting frameshifting. A prior study had suggested the presence of a frameshift attenuator element located immediately 5′ of the SARS-CoV −1 PRF signal (52). This conclusion was based on the observation that apparent frameshifting efficiency was reduced by ∼40% when an additional 150 nucleotides of sequence located 5′ of the slippery site of the SARS −1 PRF signal was included in frameshift reporter constructs. The inclusion of this sequence upstream of the IBV frameshift signal resulted in a similar reduction in −1 PRF, leading the authors to conclude that the function of this element is to specifically attenuate frameshifting. An mfold analysis revealed that this sequence may assume very stable secondary structures, leading the authors to suggest that RNA-RNA interactions between the attenuator sequence and the −1 PRF signal promote decreased rates of −1 PRF. An alternative interpretation is that the stable RNA structure assumed by this sequence simply inhibits ribosome processivity, causing a significant fraction of ribosomes to dissociate from the mRNA before encountering the −1 PRF signal. This would result in an apparent, but not actual, change in −1 PRF efficiency.

If such an element is functional, it should be conserved. In an initial survey, windows of 150 nt upstream of the −1 PRF signals from all 32 coronaviruses sequenced to date were extracted and compared. Multiple-alignment analyses using Clustal W2 (35) at both the peptide and nucleotide levels revealed that although there was good conservation of sequences within the different subgroups of viruses, there was no conservation of either primary nucleotide or peptide sequences between coronavirus subgroups (data not shown). However, since primary sequence information is not informative with regard to potential secondary- or tertiary-structure interactions, all of these sequences were folded in silico using mfold (36, 60) to address this issue. This analysis revealed that they all had the potential to form stable secondary structures, although there was no predicted consensus structure tying the different groups together. Figure 4 shows folding solutions from six representative viruses. Thus, it is possible that coronaviruses have evolved strong RNA structures to modulate frameshifting.

To determine whether the effect of the SARS-CoV attenuator element on −1 PRF is direct or indirect, a series of dual luciferase reporters was constructed either with or without the attenuator element and/or the −1 PRF signal (slippery site plus pseudoknot) (Fig. 5A). The SARS-CoV attenuator sequence, slippery site, and pseudoknot (nucleotides 13224 to 13477) were cloned between the two Renilla and firefly luciferase genes into p2luci to create the test plasmid (T+). The insertion of an additional adenosine immediately 5′ of the slippery site created a readthrough control plasmid (C+), enabling the translation of the downstream firefly gene without frameshifting. Test and control constructs lacking the attenuator sequence and containing the −1 PRF signal (pJD502 and pJD464) were described previously (47) and were used to control for the absence of the attenuator sequence. These are designated T− and C−. pLuci provided the readthrough control (RT) lacking both the attenuator and −1 PRF signal.

The apparent level of frameshifting promoted by the attenuator-containing construct (T+/C+) was 12.1%, i.e., ∼34% less than the 18.4% frameshifting promoted from the construct lacking this sequence (T+/C−) (Fig. 5B). This result compares favorably with results of the prior study (52). However, the comparison of firefly/Renilla luciferase ratios among the readthrough controls revealed that both the attenuator and −1 PRF signal inhibited ribosome processivity (Fig. 5C). Specifically, the addition of the slippery site plus pseudoknot reduced the firefly/Renilla luciferase ratios to 92% of the control plasmid lacking any inserts (compare RT/RT to C−/RT), i.e., the pseudoknot structure inhibited ribosome processivity by ∼8%. The addition of the attenuator reduced this ratio further to 62% of the no-insert readthrough control (C+/RT), i.e., the
combination of the attenuator and pseudoknot inhibited ribosome processivity by 38%. This number is nearly identical to the apparent 34% reduction in \( \frac{H11002}{F} \) PRF in the presence of the attenuator. Thus, the attenuator element does not actually inhibit \( \frac{H11002}{F} \) PRF. Rather, we suggest that this element functions to reduce the fraction of ribosomes that encounter the \( \frac{H11002}{F} \) PRF signal, from which point they may shift the reading frame and translate ORF1b. In sum, we conclude that this element has evolved as an additional means to control the stoichiometric ratio of pp1a to pp1b.

If the attenuator functions to block elongating ribosomes, then its presence should result in the accumulation of truncated peptide products. To test this, PCR products containing T7 RNA polymerase transcription promoters were synthesized, T7 RNA polymerase runoff transcripts were synthesized in vitro, \([26S]\) methionine-labeled peptides were translated in vitro, and the products were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The frameshift-promoting (T− and T−2) template lacking the attenuator primarily produced a product of 37 kDa (T− and T−2 in Fig. 5D), which is consistent with the presence of a 0-frame termination codon. The addition of the attenuator (T+) increased the size of this product to 43 kDa. Importantly, the presence of the attenuator also resulted in the production of a significant fraction of truncated peptides (indicated by an asterisk in Fig. 5D). A similar range of truncated products was produced from the readthrough reporter containing the attenuator (C+), and these were significantly less well represented in the attenuatorless readthrough control (C−). In addition, the presence of truncated products in the 37-kDa range in the C− sample is consistent with pseudoknot-induced ribosome dissociation. These results support the hypothesis that both the attenuator and the frameshift-stimulating pseudoknot cause a significant fraction of ribosomes to dissociate from the mRNA.

**DISCUSSION**

Maintaining the correct levels of coronavirus frameshifting efficiency is essential for viral infectivity. The importance of maintaining precise ratios of Gag to Gag-pol has been demonstrated for the L-A dsRNA totivirus, the HIV and MMTV retroviruses, and the 

**FIG. 5.** Apparent effect of the attenuator sequence on frameshifting is due to its affect on ribosome processivity. (A) pLuci (no-insert readthrough, i.e., RT) was employed as the parent plasmid for all subsequent constructs. The SARS-CoV frameshift signal, including the slippery site (ss) and pseudoknot (Ψk), were cloned between the two luciferase genes (test construct T−). A readthrough plasmid was created by the addition of one nucleotide (n) upstream of the slippery site (control construct C−). Similar constructs containing the upstream attenuator sequence (att) also were made (T+ and C+). The reading frame of the firefly luciferase in each of the constructs is indicated. (B) Apparent frameshifting efficiencies were assayed in Vero cells transfected with each of the indicated constructs. Firefly activity was normalized to Renilla activity, and comparisons were made between the plasmids. The test constructs were compared to the control constructs to determine frameshifting efficiency. (C) Firefly/Renilla luciferase ratios generated by the readthrough constructs containing either the pseudoknot alone (C−) or the pseudoknot and the attenuator sequence (C+) were compared to those generated from pLuci (RT). (D) Full-length luciferase proteins produced using separate transcription and translation reactions were separated by SDS-PAGE. Two different dilutions of T− (i.e., T− and T−2) are shown. An increase in the abundance of smaller proteins from attenuator-containing transcripts is indicated with an asterisk. The expected sizes of the Renilla protein with or without the attenuator at 43 and 37 kDa, respectively, are marked on the gel. The frameshift or readthrough products (Firefly/Renilla) with or without the attenuator are predicted to be 101 and 107 kDa, respectively.
which the coronavirus frameshift signals are found are very large and encode many different proteins, many of whose functions are uncharacterized. Altering the ORF1a/ORF1b protein ratios by reducing −1 PRF efficiency should result in fewer enzymes encoded by ORF1b compared to the level of proteases encoded by ORF1a.

The reduction of −1 PRF to 0.15% completely abrogated the production of infectious viral particles, whereas small amounts of infectious virus were produced at −1 PRF levels of 2.3%. These observations demonstrate that there is a lower limit of frameshifting below which coronaviruses cannot replicate to detectable levels. As anticipated, gRNA production was greatly reduced in the two viable −1 PRF mutants. After normalizing for virus levels in infected cells, the infectivity rates of the two viable slippery-site mutants remained at least 3.5 orders of magnitude less than that of the wild-type controls. One possible explanation for this dramatic effect consequent to mere 3- to 7-fold changes in −1 PRF is that decreased −1 PRF results in the decreased synthesis of ORF1ab-encoded replicase proteins, which are required for the replication of the virus. This species in turn serves as the enzyme for the synthesis of gRNA and sgRNA. In positive-sense ssRNA viruses, gRNA synthesis is greatly amplified from the negative strand, and thus the observation that the mutants decreased infectivity by >3.5 orders of magnitude, and gRNA synthesis by ∼3 orders of magnitude, may be accounted for by this amplification process. Alternatively, the decreased viral amplification may be due to a decrease in the abundance of the other nsp11 to nsp16 proteins relative to that of the nsp1 to nsp10 proteins. It is not clear from our analysis whether the loss in infectivity is directly the result of the altered RNA synthesis, protein levels, or a subsequent defect in viral particle production. Further study is needed to investigate these possibilities.

Interestingly, the slippery-site mutants had less effects on the accumulation of sgRNA (Fig. 2). Viral RNA synthesis occurs in the cytoplasm and requires only the ORF1a- and ORF1ab-encoded proteins (1, 55). The RNA-dependent RNA polymerase (RDRP) is responsible for the synthesis of both gRNAs and sgRNAs, and the ratio of these two species normally remains constant throughout the infectious cycle (50). The differences in this ratio observed between the wild-type virus and frameshift mutants demonstrate that alterations in the −1 PRF signal impact the utilization of the template RNA for the synthesis of these two RNA species. It has been shown that additional nsp12 and nucleocapsid proteins each enhance replication (41), as well as the presence of additional nsp3, a papainlike cysteine protease (50), suggesting that the processing of the polyprotein(s) is important for replication. However, those studies did not distinguish between the production of gRNA and sgRNA species. In the current study, reduced −1 PRF is expected to result in less nsp11 to nsp16 relative to the amount of nsp1 to nsp10 produced during infection. We observed a change in the ratio of gRNA and sgRNA when RNA from infected cells were analyzed by quantitative PCR. The link between these two points is, at present, unclear but likely is due to the inherent function of the nsp proteins. It is formally possible, however, that the defect in gRNA synthesis observed in the current study is due to changes in the RNA slippery-site sequence itself.

Coronavirus frameshift-promoting mRNA pseudoknot structures have evolved to fine-tune −1 PRF toward producing a golden mean of ORF1a- and ORF1b-encoded peptides. While the structural analyses presented here show that both the SARS-CoV and MHV mRNA pseudoknots are three-stemmed structures, their primary sequences and basic structural elements are significantly different from one another. Specifically, although their stem 1 elements are highly stable, the long loop 1, short stem 2, and long loop 3 of MHV are less compact than those of their analogous SARS-CoV elements. Conversely, reverse relationships apply to the stem 3 and loop 2 elements. Stabilizing/compacting any of these elements promoted increased frameshifting efficiencies, which is consistent with the notion that highly stable mRNA secondary structures cause elongating ribosomes to pause longer over the slippery site. The critical finding here is that, despite having significantly different structures, both the MHV and SARS-CoV pseudoknots promoted equivalent rates of −1 PRF. Our results show that although a class of mRNA pseudoknots that promote very high levels of frameshifting can exist, the fact that all naturally occurring coronavirus frameshift signals assayed to date promote −1 PRF efficiencies in the 10 to 20% range (46) provides indirect support for the notion that high levels of frameshifting also are incompatible with coronavirus replication. This is consistent with the golden mean model of frameshifting, i.e., it appears that each virus has evolved the right balance between more and less stable structural elements to produce the optimum rates of −1 PRF.

Coronaviruses also use translational attenuation to obtain the correct ratios of upstream and downstream viral polypeptides. The description of a sequence with a novel function, the attenuation of −1 PRF by a novel cis-acting element (52), warranted further inspection. Prior to that report, no other cis-acting elements affecting −1 PRF other than the slippery site and downstream stimulatory element had been described, and no concrete model for how the attenuator functioned was proposed. The presence of a large number of previously unavailable coronavirus genomes allowed an initial phylogenetic analysis. The alignment of these sequences revealed a large degree of diversity among the analogous regions in both homology and length; the smallest coronavirus genomes contained the shortest analogous sequences, and the most diverse genomes showed the least homology. This indicated that no specific attenuator sequence is conserved among coronaviruses, suggesting that the element does not actually directly affect the process of frameshifting. However, the computational prediction of highly stable secondary structures engendered the hypothesis that the attenuator element impedes the processivity of elongating ribosomes on the viral mRNA, causing them to dissociate prior to encountering the −1 PRF signal, i.e., a translational attenuation model. As supported by the experiments in the current study, the addition of the attenuator sequence upstream of the dual luciferase assay confirmed that its presence resulted in an apparent decrease in −1 PRF. However, the critical experiment, comparing the readthrough controls, revealed that while the −1 PRF signal itself promoted a small but significant decrease in the production of the full-length product (presumably due to the presence of the mRNA pseudoknot), the addition of the attenuator decreased this value by nearly 40%. Further, in vitro translation assays
demonstrated that the attenuator sequence promoted the increased synthesis of prematurely terminated peptide products. Again, although indirect, this provides indirect support of the golden mean hypothesis, in that they show that coronaviruses have evolved a second cis-acting element to limit the fraction of ribosomes able to eventually translate ORF1b. In sum, the data presented here support a model in which a combination of ribosomes able to eventually translate ORF1b. In sum, the increased synthesis of prematurely terminated peptide products. Again, although indirect, this provides indirect support of the frameshifting efficiency serve translation elongation factor I alpha differentially affect programmed ribosomal frameshifting and drug sensitivity. RNA 38:70–88.


from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. J. Virol. 65:5111–5117.


