THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 49, pp. 37952–37961, December 8, 2006

The SL1–SL2 (Stem-Loop) Domain Is the Primary Determinant for Stability of the Gamma Retroviral Genomic RNA Dimer*

Received for publication, August 3, 2006, and in revised form, September 8, 2006 Published, JBC Papers in Press, September 19, 2006, DOI 10.1074/jbc.M607380200

Cristina Gherghe and Kevin M. Weeks[1]

From the Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290

Retroviral genomes are assembled from two sense-strand RNAs by noncovalent interactions at their 5′ ends, forming a dimer. The RNA dimerization domain is a potential target for antiretroviral therapy and represents a compelling RNA folding problem. The fundamental dimerization unit for the Moloney murine sarcoma gamma retrovirus spans a 170-nucleotide minimal dimerization active sequence. In the dimer, two self-complementary sequences, PAL1 and PAL2, form intermolecular duplexes, and an SL1–SL2 (stem-loop) domain forms loop-loop base pairs, mediated by GACG tetraloops, and extensive tertiary interactions. To develop a framework for assembly of the retroviral RNA dimer, we quantified the stability of and established nucleotide resolution secondary structure models for sequence variants in which each motif was compromised. Base pairing and tertiary interactions between SL1–SL2 domains contribute a large free energy increment of −10 kcal/mol. In contrast, even though the PAL1 and PAL2 intermolecular duplexes span 10 and 16 bp in the dimer, respectively, they contribute only −2.5 kcal/mol to stability, roughly equal to a single new base pair. First, these results emphasize that the energetic costs for disrupting interactions in the monomer state nearly balance the PAL1 and PAL2 base pairing interactions that form in the dimer. Second, intermolecular duplex formation plays a biological role distinct from simply stabilizing the structure of the retroviral genomic RNA dimer.

Retroviruses carry their genetic information in the form of RNA and replicate through a DNA intermediate (1). The retroviral genome consists of two noncovalently linked sense-strand RNAs, forming a dimer (2–4). Formation of the genomic RNA dimer appears to be important for key stages of the retroviral replication cycle, including selective packaging of genomic RNA (4–6), reverse transcription, and recombination (7–10).

Although a retroviral RNA genome forms multiple contacts between its constituent monomers, the most stable point of contact is situated at the 5′ end of the RNA (11–14). Our laboratory has defined and structurally characterized a minimal dimerization active sequence (MiDAS)[2] of 170 nucleotides for a representative gamma retrovirus, the Moloney murine sarcoma virus (MuSV), as the major 5′ dimerization motif (15–17) (Fig. 1). The MiDAS is a structurally independent domain that dimerizes efficiently to yield homogeneous monomer and dimer populations *in vitro*. The MiDAS RNA dimerizes under biologically relevant conditions of both temperature and ionic environment, in the absence of proteins. Moreover, RNAs spanning sequences closely related to the MiDAS are sufficient to mediate RNA dimerization and encapsidation *in vivo* when inserted into nonviral mRNAs (3, 18). Thus, the MiDAS domain is an important system for studying retroviral RNA dimerization as it contains all of the RNA sequence elements required for the high selectivity and affinity of this process both *in vitro* and *in vivo*.

Structures of the MiDAS domain for MuSV in the monomeric starting state and in the final dimer state have been mapped at single nucleotide resolution using selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) (15, 17). Taken together with extensive prior work (19–27), these structural studies emphasize that there are three major components of the retroviral dimer: two self-complementary (palindromic) sequences, termed PAL1 and PAL2, plus a double stem-loop (SL1–SL2) domain (Fig. 1). The two stem-loops, SL1 and SL2, contain conserved GACG tetraloops (19) and function together as an independent domain (16, 25, 26, 28).

The existence of self-complementary sequences with the potential to form intermolecular duplexes is broadly conserved among retroviral RNA genomes (19, 22). Similarly, stable stem-loop motifs, terminating in GACG tetraloops, are prominent features for at least two retroviral genera, the gamma retroviruses (19) and the nonprimate lentiviruses (29). Understanding the relative structural contributions of self-complementary and GACG stem-loop domains is essential to develop models for the broad organizing principles of RNA genome structure and the ability of retroviruses to function as simple diploid entities in biology.

Significant conformational changes occur in each of the three PAL1, PAL2, and SL1–SL2 domain structural building blocks during dimerization (15–17, 20–23) (Fig. 1). PAL1 resides in a stem-loop structure in the monomer state and forms an extended intermolecular duplex involving both genomic RNA strands in the dimer (Fig. 1, *green sequences*). PAL2 is relatively unstructured in the MiDAS monomer and lies in a flexible domain that is connected to the rest of the RNA via a stable anchoring helix (Fig. 1*A*, *shaded box*) (15). Upon dimerization, the unstructured PAL2 sequences from two monomers also form an extended intermolecular duplex (Fig. 1, *red sequences*). As the PAL2 intermolecular duplex forms, the

* This work was supported by National Institutes of Health Grant GM064803 (to K. M. W. and Andrew Kaplan). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed. E-mail: weeks@unc.edu.

[2] The abbreviations used are: MiDAS, minimal dimerization active sequence; MuSV, Moloney murine sarcoma virus; SHAPE, selective 2′-hydroxyl acylation analyzed by primer extension; NMIA, *N*-methyl isatoic anhydride.

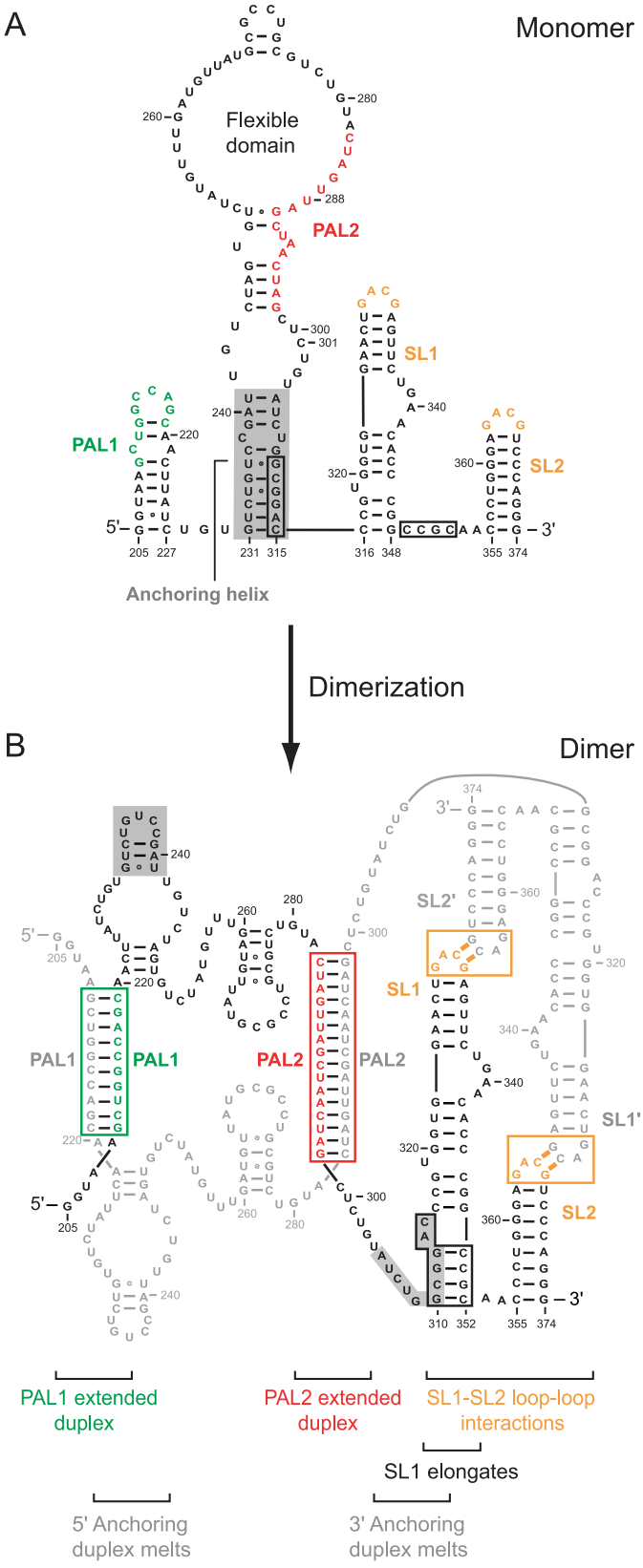


FIGURE 1. **Secondary structures and conformational changes that accompany dimerization of the MiDAS domain for the MuSV gamma retrovirus.** Major interaction sites (PAL1, PAL2, and the SL1–SL2 tetraloops) are emphasized in color. The anchoring helix and the nucleotides that yield a 4-bp extension of SL1 are emphasized with *shaded* and *black boxes*, respectively. Multiple conventions are in use for the PAL1, PAL2, SL1, and SL2 elements; we use the nomenclature originally introduced for this system (19).

anchoring helix characteristic of the monomer state melts (Fig. 1*B, shaded boxes*) (17).

At the 3′ end of the MiDAS, SL1 and SL2 form cross-strand loop-loop interactions involving conventional G-C base pairs and additional noncanonical interactions (26) with SL2′ and SL1′ sequences, respectively, from the second RNA strand (Fig. 1, *orange sequences*). During dimerization, SL1 also extends by 4 bp (Fig. 1*B, black box*) (16). When this conformational change occurs, the SL1–SL2 domain forms a compactly folded and high affinity tertiary structure domain (16).

Formation of the PAL2 intermolecular duplex and extension of SL1 in the dimer state both require disruption of the anchoring helix. Thus, PAL2 and SL1–SL2 domain interactions are potentially coupled during dimerization.

We analyze the contribution of each of these key structural features to the stability of the MuSV genomic RNA dimer. In the context of the MiDAS domain, the SL1–SL2 domain interaction makes the largest contribution to overall thermodynamic stability in the dimer. PAL1 and PAL2 form 10- and 16-bp duplexes in the dimer and, as isolated structural elements, would be expected to have large thermodynamic stabilities of up to −15 kcal/mol. Surprisingly, the net thermodynamic increment for forming either the PAL1 or PAL2 intermolecular duplex in the dimer is small. We infer that the free energy released upon formation of either extended duplex in the dimer is almost exactly balanced by the energetic cost to break pre-existing interactions in the starting monomer structure. These data indicate that the highly conserved ability to form intermolecular duplexes in retroviral RNA genomes may be more important for retroviral processes such as modulating protein-binding sites or regulating dimerization kinetics than simply stabilizing the dimer state.

EXPERIMENTAL PROCEDURES

Retroviral RNA Constructs—DNA templates for *in vitro* transcription of the native and mutant MiDAS RNA constructs (PAL1Stb, PAL2Del, Stb-Del, AA/AA) were generated by PCR from the pLNBS (30) plasmid or from synthetic oligonucleotide (Midland) templates. RNA constructs were generated by T7 RNA polymerase-mediated transcription (200 μl, 37 °C, 6 h; containing 80 mM HEPES (pH 7.5), 40 mM dithiothreitol, 0.01% (v/v) Triton X-100, 2 mM spermidine, 20 mM $MgCl_2$, 2 mM each nucleoside triphosphate, ~10 μg of PCR-generated template, 200 units of SUPERase-In (Ambion), and 0.07 mg/ml polymerase). Internally labeled RNAs were generated using 25 μCi of [α-^{32}P]ATP and 0.5 mM unlabeled ATP. RNAs were purified by denaturing gel electrophoresis (8% (w/v) polyacrylamide, 7 M urea), excised from the gel, eluted overnight (into 0.5 M sodium acetate (pH 6.0), 1 mM EDTA; 4 °C), concentrated by ethanol precipitation, and stored in TE (10 mM Tris (pH 7.5), 1 mM EDTA) at −20 °C.

Equilibrium Dimerization Affinities—All dimerization reactions were performed at 37 °C and under an approximately physiological ion environment (200 mM potassium acetate, 5 mM $MgCl_2$, pH 7.5). Internally ^{32}P-labeled RNA (0.05 nM) was combined with the identical unlabeled RNA (at 0.1–300 nM) in 6.4 μl. RNAs were denatured at 90 °C, snap-cooled on ice for 3 min, treated with 3.2 μl of 5× dimer buffer (250 mM HEPES

(pH 7.5), 1 M potassium acetate (pH 7.5), 25 mM $MgCl_2$) (15) and 6.4 μl of loading dye (30% glycerol, 0.01% xylene cyanol, and bromphenol blue), and incubated at 37 °C for either 2 or 8 h. Monomer and dimer forms were resolved on nondenaturing gels (8% polyacrylamide, 1× TBE) run at a gel temperature of 25 °C (1 h at 7 watts). RNAs were also resolved in gel and running buffers containing 5 mM $MgCl_2$. Gels were pre-run for 20 min prior to loading, and the running buffer was re-equilibrated every 10 min to maintain uniform ion concentrations. Monomer and dimer species were visualized by phosphorimaging, and the dimerization dissociation constant (K_{dimer}) was obtained by fitting to Equation 1,

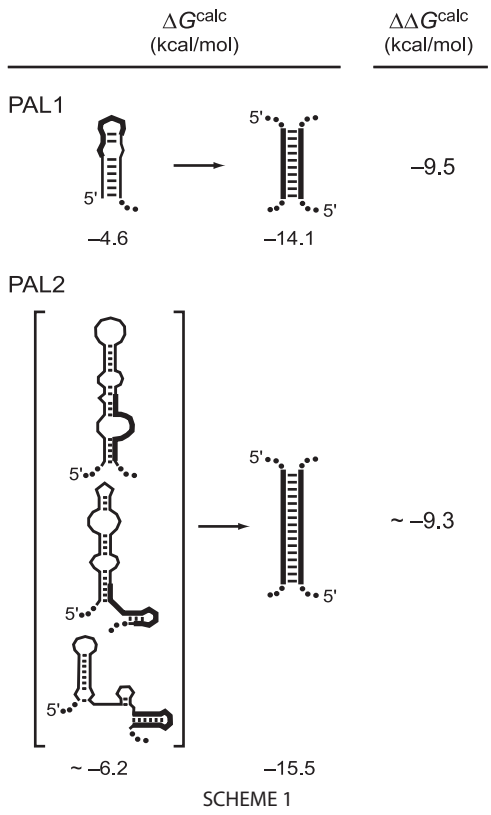
$$\frac{A}{4C_T}\left((K_{dimer} + 4C_T) - \sqrt{K_{dimer}^2 + 8K_{dimer}C_T}\right) \qquad \text{(Eq. 1)}$$

where A is the fraction dimer at saturating RNA concentrations and C_T is the total concentration of RNA.

SHAPE Analysis of Native and Mutant Dimers—SHAPE experiments were performed on native or mutant MiDAS RNAs that contained a 3′ nonviral RNA cassette that provides an efficient DNA primer binding site (31), linked to the MiDAS RNA via an AAAU sequence. Control experiments showed that introduction of the 3′ cassette had no effect on dimerization affinities. The MiDAS RNA construct (9 pmol) was heated at 90 °C for 3 min in 2.7 μl of water, cooled on ice, and treated with 1.8 μl of 5× dimerization buffer at 37 °C for 2 h. The RNA solution was then treated with 1 μl of NMIA (100 mM in anhydrous DMSO) and allowed to react for 45 min (~4 half-lives) (31, 32).

Primer Extension—DNA primers were complementary to the 3′ end of the RNA structure cassette (5′-GAA CCG GAC CGA AGC CCG-3′) or to SL1 (5′-CAG AAC TCG TCA GTT CCA CCA-3′) (15). NMIA-modified RNA (2 μl, 1.8 pmol) was annealed to a 5′-^{32}P-labeled DNA primer (1 μl, 1 pmol) in 12 μl total volume and incubated at 60 (6 min) and 35 °C (10 min). Reverse transcription buffer (7 μl; 143 mM Tris acetate (pH 8.4), 214 mM potassium acetate, 7.14 mM $MgCl_2$, 1.43 mM each dNTP, 14.3 mM dithiothreitol) was added, and subsequent primer extension steps were performed as described (15, 33), using ThermoScript reverse transcriptase (Invitrogen) at 50 °C. cDNA fragments were resolved on a series of 14% (w/v) polyacrylamide gels to achieve nucleotide resolution throughout the analyzed region. Band intensities were quantified using semi-automated footprinting analysis software (34). SHAPE reactivity data for the mutants were normalized to that of the native MiDAS using nucleotide 314 for PAL1Stb and nucleotide 319 for the PAL2Del and Stb-Del mutants.

Free Energy Calculations—RNAstructure (35) was used to estimate the change in free energy associated with formation of the PAL1 and PAL2 intermolecular duplexes relative to the structures these sequences form in the monomeric initial state (illustrated in Scheme 1). Structures for PAL1 and PAL2 in the monomer state were obtained using strong and moderate SHAPE reactivities to constrain single-stranded nucleotides (15). PAL1 stability was calculated for nucleotides 205–227. PAL2 stability was estimated for nucleotides 244–299, as the mean folding free energy for three postulated (15) interconverting structures in the 231–315 domain and subtracting the free energy of the anchoring duplex (nucleotides 231–243/300–315). Stabilities for the PAL1 and PAL2 intermolecular duplexes in the dimer state were estimated using the bimolecular folding mode in RNAstructure (35).



SCHEME 1

Incremental Dissociation Constants for Individual Dimerization Steps—For the dimerization step involving SL1–SL2 domain interactions, the dimerization constant K_1 was measured directly as the affinity of the Stb-Del mutant. For steps involving SL1–SL2 interactions plus either PAL1 or PAL2 formation, K_2 and K_2' were calculated from K_{dimer} for the PAL2Del and PAL1Stb mutants, respectively, as $K_{dimer} = K_1K_2$. For the third step, K_3 and K_3' were calculated from K_{dimer} for dimerization of the native MiDAS, $K_{dimer} = K_1K_2K_3$. Free energies for each dimerization step were calculated as $\Delta G = -RT \ln(1/K_x)$ at 37 °C.

RESULTS

Strategy—To quantify the energetic contribution of each element of the MiDAS RNA to stability of the retroviral dimer, we constructed mutants that compromise precisely each of the three primary contributors to dimerization: PAL1, PAL2, and the SL1–SL2 domain (Fig. 2). The ability of PAL1 to form an intermolecular duplex in the dimer was abolished by stabilizing the PAL1 stem-loop in its monomer form to make the PAL1*Stb* mutant (Fig. 2*A*). PAL2 was compromised by a 12-nucleotide deletion in the PAL2*Del* mutant (Fig. 2*B, dashed line*). This deletion completely abrogates PAL2 intermolecular duplex formation but does not cause significant structural changes in other parts of the MiDAS RNA monomer (15). We also constructed a PAL1*Stb*-PAL2*Del* double mutant, *Stb-Del*, that

eliminates both PAL1 and PAL2 duplexes in the dimer (Fig. 2*C*).

Two SL1–SL2 domains form an extensive tertiary interface mediated by loop-loop interactions involving the GACG sequences in the apical loops (26) and by interdigitated interactions between closely packed U-shaped SL1–SL2 motifs (16). SL1 from one RNA interacts with SL2′ from the second strand (16). No simple, compact mutation can completely disrupt this tertiary interface; however, we selectively disrupted the loop-loop component of the SL1–SL2 domain tertiary structure by converting the conserved GACG sequence in the apical tetraloops to GAAA (Fig. 2*D*, *AA/AA mutant*). GAAA sequences form a stable local loop structure but cannot form strong cross-strand loop-loop interactions. The double loop mutant also prevents noncognate SL1-SL1′ or SL2-SL2′ interactions from forming (16).

Dimerization Affinities for Native and Mutant MiDAS RNAs—An important feature of the MiDAS domain is that dimerization does not require extreme temperature or solution conditions. Dimerization proceeds efficiently at 37 °C and at roughly physiological ionic strength (200 mM potassium acetate, 5 mM $MgCl_2$ (pH 7.5)). This well behaved dimerization activity appears to reflect that inhibitory RNA sequences are absent in the MiDAS domain. Both native and mutant RNAs form homogeneous monomer and dimer species as judged by nondenaturing gel electrophoresis (Fig. 3).

Exploratory work showed that RNA dimers in which either the PAL1 or PAL2 intermolecular duplex has formed can be detected by nondenaturing gel electrophoresis, both in the presence and absence of Mg^{2+}. Dimerization at the SL1–SL2 domain specifically requires the presence of Mg^{2+} in the gel and running buffer in order to be detected during electrophoresis. Thus, dimerization for all RNA constructs was visualized in both Mg^{2+}-free and Mg^{2+}-containing gels (Fig. 3, *A* and *B*). Dimerization reactions were performed as a function of RNA concentration, and apparent dimerization dissociation constants (K_{dimer}) decrease by 3-fold or less when the equilibration period is extended from 2 to 8 h. We therefore take the 8-h value as a good approximation for fully equilibrated retroviral RNA dimers.

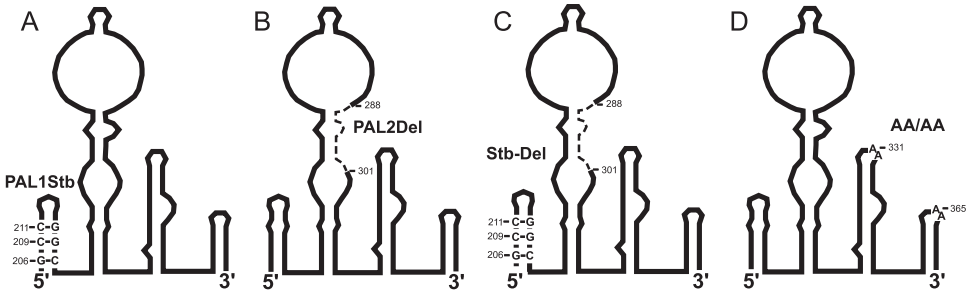


FIGURE 2. **Secondary structure representations of MiDAS variants in the monomer state.** Mutations designed to selectively disrupt formation of secondary structure elements in the final dimer are shown explicitly.

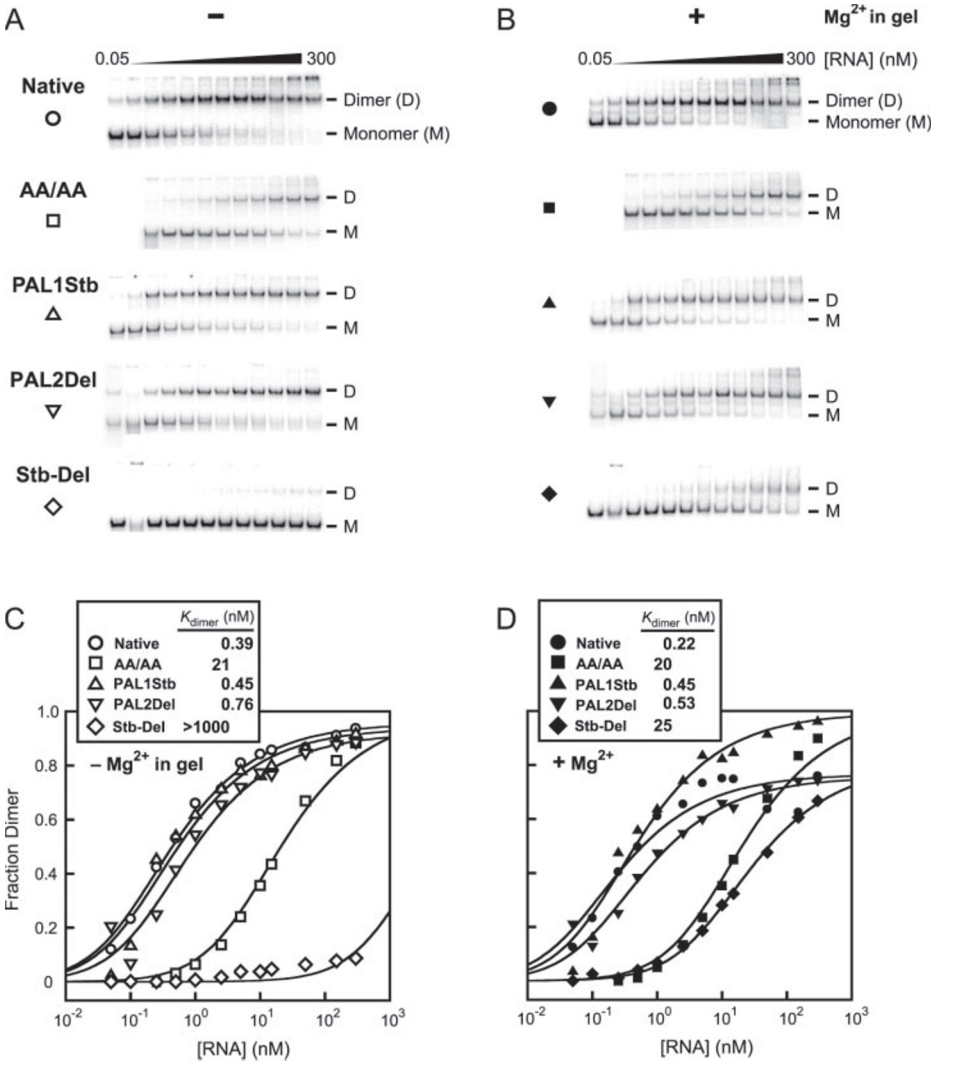


FIGURE 3. **Dimerization affinities for native and mutant MiDAS RNAs.** *A* and *B*, dimer formation visualized by nondenaturing gel electrophoresis, monitored in the absence (−) and presence (+) of 5 mM Mg^{2+}. RNA concentrations were 0.05, 0.1, 0.25, 0.5, 1, 2.5, 5, 10, 15, 50, 150, and 300 nM; RNAs were equilibrated for 8 h prior to resolution by gel electrophoresis. *C* and *D*, quantitative equilibrium dissociation curves for dimerization.

The native MiDAS RNA forms a very high affinity dimer, independent of the presence of Mg^{2+} in the gel. K_{dimer} values are 0.39 and 0.22 nM, as visualized in gels run in the absence and presence of Mg^{2+}, respectively (Fig. 3, *C* and *D*, *open* and *closed circles*).

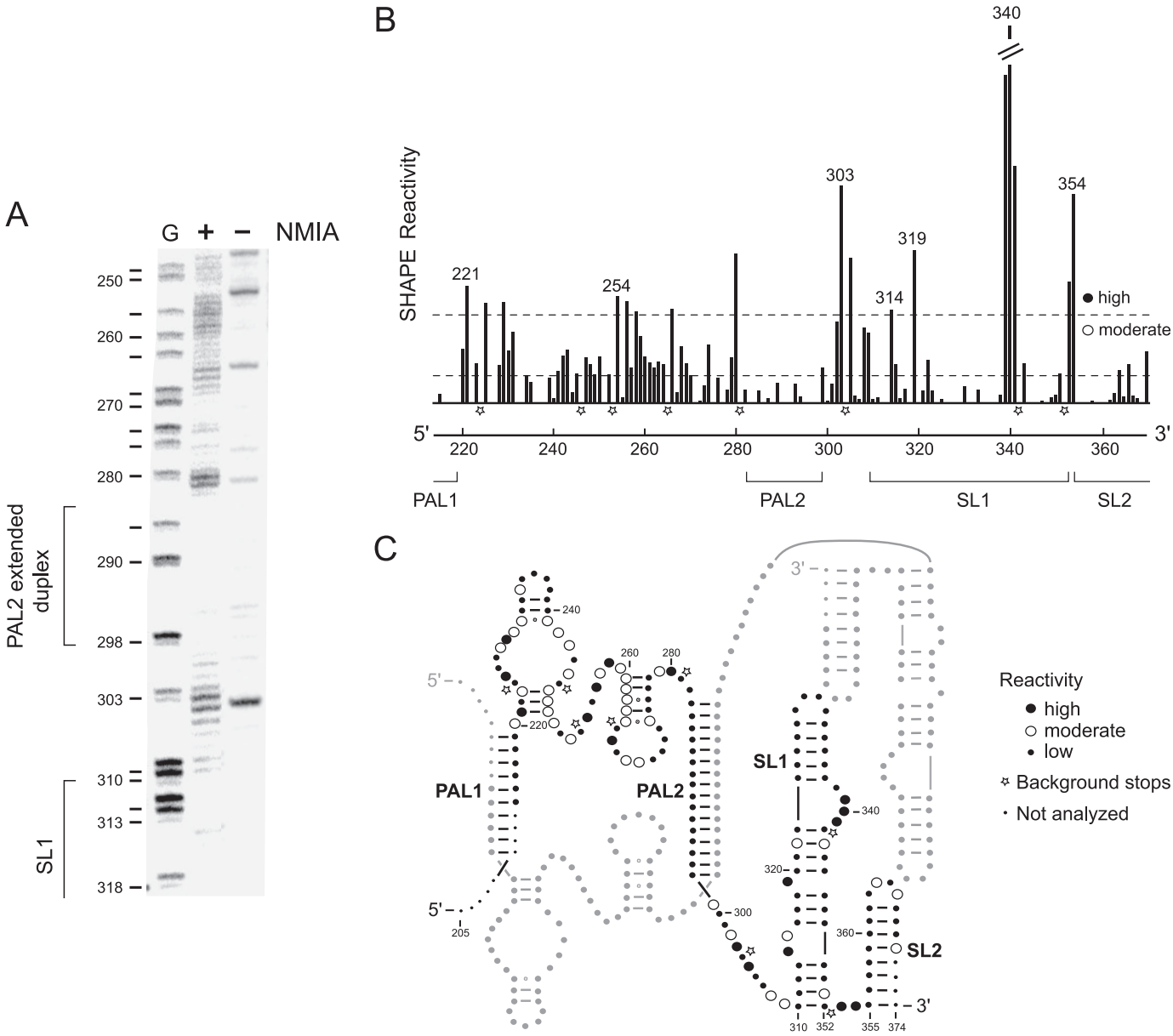


FIGURE 4. **SHAPE analysis of the native MiDAS RNA in the dimer state.** *A*, primer extension reactions resolved on a sequencing gel for RNAs treated in the presence (+) or absence (−) of the structure-selective reagent, NMIA. *G*, guanosine sequencing lane; sequence is offset by one nucleotide relative to NMIA lanes, *numbering* refers to NMIA lanes. *B*, histogram of absolute SHAPE reactivities, minus background, as a function of nucleotide position. *C*, superposition of SHAPE information on a secondary structure model for the MiDAS RNA dimer. For clarity, only one of the two strands in the dimer is annotated.

We then evaluated the thermodynamic contribution of the SL1–SL2 domain to dimer formation using the AA/AA mutant in which native loop-loop interactions are disrupted. Dimerization affinities are similar, independent of whether Mg^{2+} is present in the gel. However, dimerization is weakened by ~100-fold relative to the native sequence (K_{dimer} = 20 nM; Fig. 3, *open* and *closed squares*). Thus, the SL1–SL2 domain is a significant contributor to the stability and structure of the dimer state.

In contrast, the PAL1Stb and PAL2Del RNA variants, which eliminate the ability to form either PAL1 or PAL2, dimerize with affinities that are within 2.5-fold of that for the wild type (Fig. 3, *triangles* and *inverted triangles*). These data indicate that formation of either the PAL1 or PAL2 intermolecular duplex makes a very small contribution to the stability of the dimer if the SL1–SL2 domain is present.

We next evaluated the dimerization activity of the Stb-Del mutant, in which both PAL1 and PAL2 are compromised. This RNA is only capable of dimerizing by forming interactions in the SL1–SL2 domain, including via cross-strand loop-loop base pairs at the GACG tetraloops (see Fig. 1). Dimer stability is strongly dependent on whether Mg^{2+} is present during nondenaturing gel electrophoresis (Fig. 3, *A* and *B, Stb-Del panels*). Formation of a dimer species is barely detectable in gels run in the absence of Mg^{2+} (apparent K_{dimer} ~ 4.2 μM), whereas stable dimer formation is readily detected in Mg^{2+}-containing gels and is characterized by a K_{dimer} of 25 nM.

These results demonstrate that two SL1–SL2 domains form a high affinity interaction that is strongly dependent on the presence of Mg^{2+}. In the presence of the stable Mg^{2+}-mediated SL1–SL2 domain interaction, PAL1 and PAL2 each then make

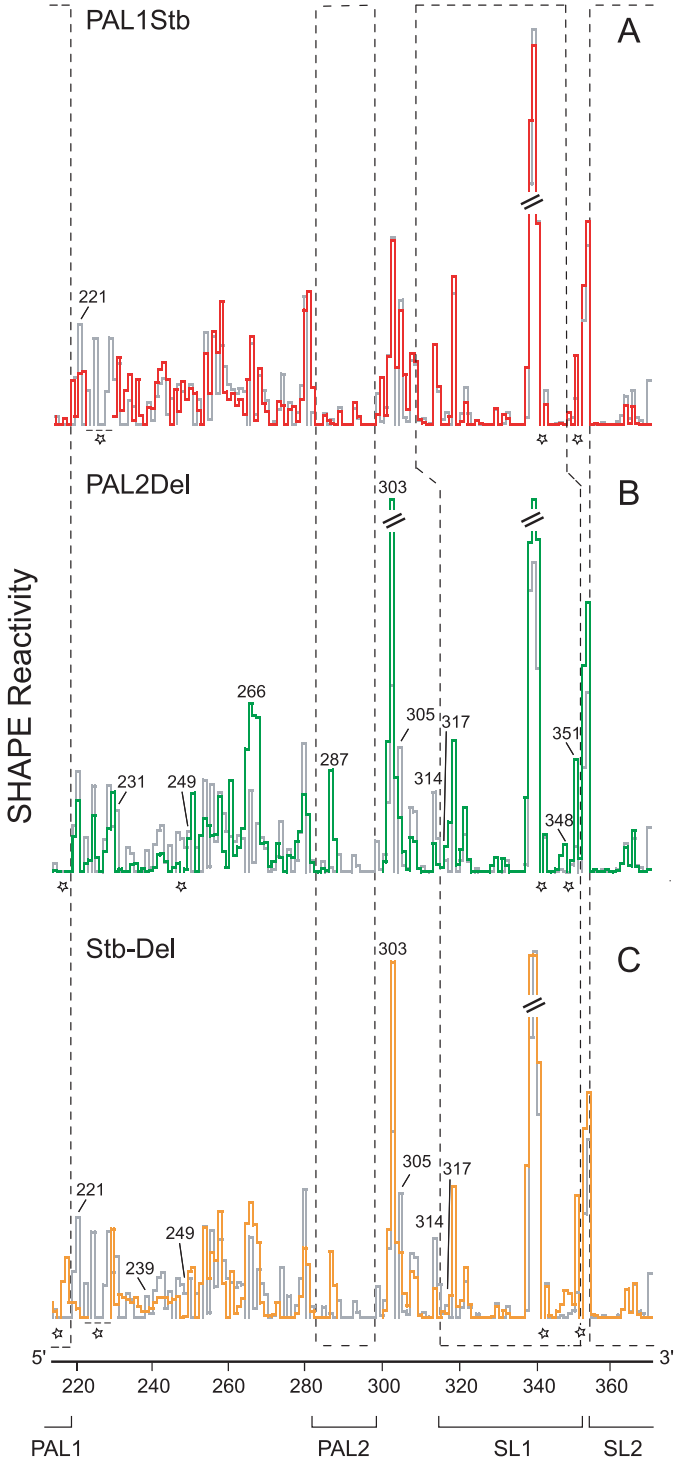


FIGURE 5. **Structural differences between native and mutant MiDAS RNAs analyzed by SHAPE chemistry.** In each histogram, mutant and native RNAs are shown by *colored* and *gray lines*, respectively. *Stars* indicate a small number of positions that could not be analyzed because of strong SHAPE chemistry-independent stops to primer extension.

small additional contributions to dimer stability.

Structural Analysis of Dimers—We explored the specific structural differences in the dimers formed by the PAL1Stb, PAL2Del, and Stb-Del mutants using single nucleotide resolution RNA SHAPE analysis (31, 32). In the SHAPE approach, an RNA is treated with a strongly electrophilic reagent, NMIA, that reacts preferentially with 2′-hydroxyl groups in conformationally flexible nucleotides. Flexible nucleotides react to form 2′-*O*-ester adducts, which are then detected as stops to reverse transcriptase-mediated primer extension.

The secondary structure of the MiDAS region in the dimer state was recently established by SHAPE (17). We confirmed the secondary structure of our MiDAS dimer by subjecting the native dimer RNA to 2′-*O*-esterification and detecting cDNA products in a sequencing gel (Fig. 4*A*). Absolute SHAPE reactivities were calculated by subtracting background band intensities observed for reactions omitting NMIA from those obtained in the presence of the reagent (Fig. 4*B*). Local nucleotide flexibility information was obtained for almost every position in the MiDAS RNA.

The nucleotides comprising the PAL1 and PAL2 intermolecular duplexes are uniformly unreactive toward SHAPE chemistry (Fig. 4*C*). SL1 and SL2 form stable stem-loop structures connected by a flexible two-nucleotide linker (nucleotides 353–354). The PAL1, PAL2, and SL1–SL2 domain structures are then connected by two highly reactive, and therefore flexible, linker regions (nucleotides 220–282 and 299–309; Fig. 4, *B* and *C*). SHAPE analysis thus strongly supports the secondary structure model for the dimer state shown in Fig. 4*C*.

SHAPE Analysis of MiDAS Mutants—Absolute nucleotide reactivities were also obtained for the PAL1Stb, PAL2Del, and Stb-Del mutants using SHAPE chemistry. Structural differences that distinguish each mutant from the native sequence in the dimer state can be evaluated in a straightforward way using reactivity histograms (Fig. 5). SHAPE reactivity for the native sequence is *gray* and reactivities for the mutants are shown in color in Fig. 5. Higher and lower reactivities for a mutant at a given position indicate that the mutant sequence is more flexible or more constrained, respectively, relative to wild type.

The reactivity profile for the PAL1Stb variant superimposes very closely with that for the wild type sequence over nearly the entire length of the MiDAS domain (Fig. 5*A*, compare *red* and *gray lines*). The most significant difference is the 2-fold reduced reactivity of nucleotide 221 in the PAL1 stem at the 5′ end of the RNA, indicating that this position is more constrained in the PAL1Stb mutant. The observed low reactivity at nucleotide 221 reflects the duplex stabilizing effect of the mutation introduced at this position (Fig. 2*A*). The primer extension reaction also records the increased stability of the PAL1 stem-loop in an indirect way, because the stabilized PAL1 structure yields a prominent stop to primer extension by the reverse transcriptase enzyme (nucleotides 223–229; Fig. 5*A*, *dashed line*). Thus, SHAPE analysis indicates that the PAL1Stb mutant folds to a dimer structure that is very similar to that of the native sequence, except that PAL1 remains in its monomeric conformation, embedded in a stem-loop (Fig. 6*A*).

The PAL2Del mutation was shown previously to induce only relatively minor changes to the structure of the monomer state. The PAL2Del is compatible with the monomer conformation because PAL2 lies in the middle of a large flexible domain (Fig. 1*A*). In contrast, the PAL2Del mutation introduces many changes to the structure of the final dimer (Fig. 5*B*, compare *green* and *gray histograms*). The most important structural differences lie in three distinct regions in the PAL2Del mutant RNA. Positions 231–249 and 305–314 span the anchoring helix

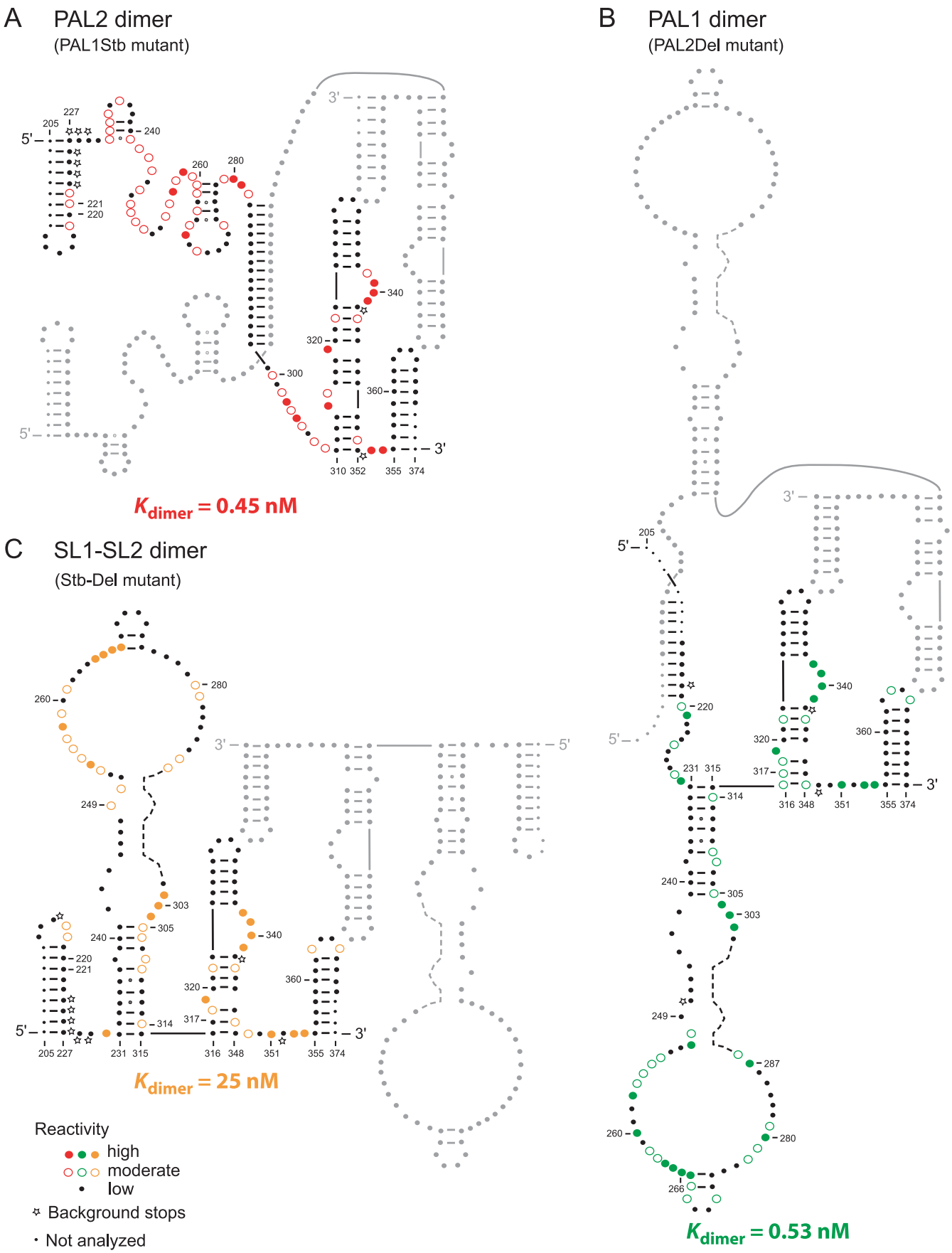


FIGURE 6. **Superposition of SHAPE information on secondary structure models for MiDAS RNA mutants in the dimer state.** K_{dimer} for each mutant was obtained from the data shown in Fig. 3*D*.

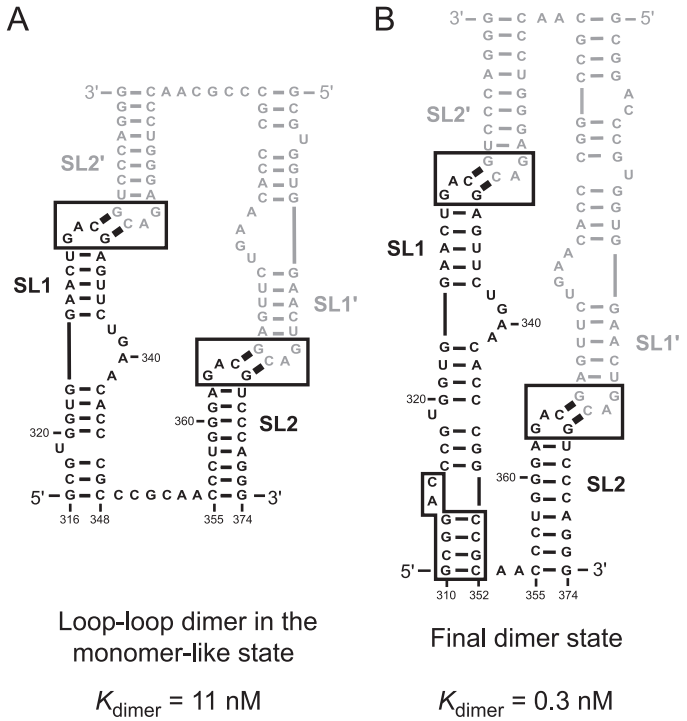


FIGURE 7. **Structures and dimerization dissociation constants for the monomer-like and dimer states of the isolated SL1–SL2 domain (16).**

and are 2–10-fold less reactive in the mutant than in the native dimer. Next, nucleotides 317–318 and 351 are 3–5-fold more reactive in the mutant, consistent with a model in which SL1 remains in the shorter, monomer-like conformation for this stem-loop (see Fig. 1*A*). Finally, nucleotides 265–268, 287–288, and 303 in the PAL2 sequence and in other parts of the flexible domain are 3–12-fold more reactive in the mutant, consistent with failure to form the PAL2 helix. Together, the comparative SHAPE reactivity information is consistent with a secondary structure model for the PAL2Del dimer in which the PAL1, but not the PAL2, intermolecular duplex forms (Fig. 6*B*). An important feature of the PAL2Del variant in the dimer state is that the anchoring helix (nucleotides 231–241/305–315) does not melt, and the SL1–SL2 domain remains in the monomer-like conformation (compare Figs. 4*C* and 6*B*).

The Stb-Del double mutant dimer has a SHAPE reactivity profile broadly consistent with a simple combination of the reactivity profiles for both the PAL1Stb and PAL2Del mutants (Fig. 5*C, orange*). Similar to the PAL1Stb mutant, the double mutant shows up to an 8-fold decreased reactivity at the 5′ end of the RNA (nucleotides 220–223), suggesting that PAL1 remains in its monomeric stem-loop conformation. Again, the stabilized PAL1 stem-loop yields a strong stop for primer extension (Fig. 5*C, dashed line* at nucleotides 224–229). The Stb-Del dimer also contains features analogous to those seen with the PAL2Del mutant. Nucleotides 239–249 and most nucleotides spanning 305–314 are unreactive, indicating that the anchoring duplex does not melt in this mutant dimer structure. Nucleotides 317 and 347–349 are more reactive toward NMIA, supporting the short monomer-like conformation for SL1. The reactivity profile of this mutant is consistent with a secondary structure model in which most of the RNA remains in a monomer-like conformation. Two Stb-Del mutant monomers therefore dimerize (see Fig. 3*B*) primarily through loop-loop contacts in the SL1–SL2 domain, without support from either the PAL1 or PAL2 intermolecular duplexes (Fig. 6*C*).

DISCUSSION

Three pieces of information allow us to define a thermodynamic framework for gamma retroviral RNA dimerization under near-physiological conditions of temperature and ionic environment. First, we have quantified the thermodynamic contribution of each of the PAL1, PAL2, and SL1–SL2 interactions to overall dimer stability. Eliminating the ability of either the PAL1 or PAL2 sequences to form intermolecular duplexes does not significantly reduce the stability of the genomic RNA dimer (Fig. 3). An independent analysis has also emphasized that PAL2 makes a small contribution to dimer stability (36). The most significant determinant of dimer stability is the SL1–SL2 domain. Disrupting the cross-strand loop-loop base pairing component of the SL1–SL2 domain reduces dimerization affinity by 100-fold. RNAs containing mutations in the SL1 and SL2 loop sequences (37, 38) are likely to have a significantly altered global architecture in the dimer state, as compared with native sequence RNAs.

Second, structural analysis of native and mutant dimers at single nucleotide resolution using RNA SHAPE chemistry indicates that folding of PAL2 and the SL1–SL2 domain are interdependent (Figs. 4–6). Both formation of the PAL2 intermolecular duplex and extension of the SL1 helix to achieve the final dimer conformation require disruption of the anchoring helix.

Third, our laboratory has previously measured dimerization affinities for the isolated SL1–SL2 domain in both monomer and dimer conformations (Fig. 7) (16). By comparing dimerization affinities for the isolated SL1–SL2 domain with those for the complete MiDAS motif, we can quantify the additional energetic increment, if any, resulting from formation of the PAL1 and PAL2 intermolecular duplexes.

We describe each step in the thermodynamic pathway for gamma retroviral RNA dimerization from two perspectives: energetically, in terms of the microscopic or incremental dimerization dissociation constant, and structurally, as a nucleotide resolution secondary structure for each dimer intermediate. Our thermodynamic model for dimerization includes the following three major steps: the high affinity and bimolecular association of two monomers through cross-strand loop-loop interactions in the SL1–SL2 domain plus two low affinity unimolecular rearrangements to form the PAL1 and PAL2 intermolecular duplexes (Fig. 8).

The SL1–SL2 Domain Significantly Stabilizes the MiDAS Dimer—Structural analysis indicates that the PAL1Stb and PAL2Del mutants precisely affect the targeted structural element and that neither mutation causes unexpected non-native interactions to form (Fig. 6). Therefore, the loop-loop interaction in the SL1–SL2 domain in the first step can be estimated to be comparable with the dimerization affinity for the Stb-Del mutant because the SL1–SL2 interaction is the only possible dimerization motif remaining in this mutant (Fig. 8, *orange labels*). The SL1–SL2 dimerization interaction is strongly modulated by Mg^{2+}. In the presence of the divalent ion, the dimerization constant is 25 nM (Fig. 3*D*). This value is similar to the dimerization constant meas-

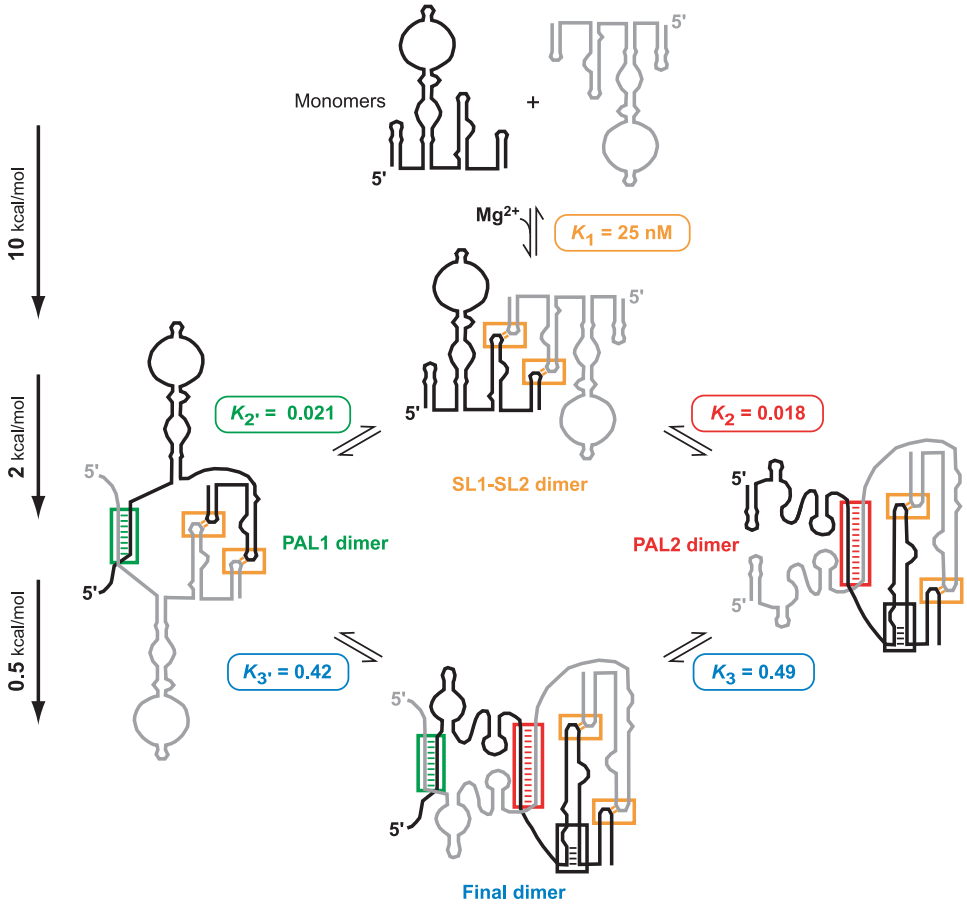


FIGURE 8. **Thermodynamic framework for MuSV genomic RNA dimerization.** Secondary structure models for the SL1–SL2, PAL1, and PAL2 dimers are extrapolated from the secondary structures of the Stb-Del, PAL2Del, and PAL1Stb dimers, respectively, as determined by SHAPE analysis. Newly formed interactions in each state are emphasized in *color* and with *boxes*. Dissociation constants (K_x) are the microscopic, or incremental, constant for each step and are calculated from the global K_{dimer} for each mutant.

ured previously for the isolated SL1–SL2 domain, also in the presence of Mg^{2+} (Fig. 7*A*) (16). Thus, either as an independent motif or in the context of the MiDAS, the SL1–SL2 domain contributes a large free energy increment to dimerization, corresponding to −10 kcal/mol (Fig. 8).

PAL1 and PAL2 Make Small Contributions to Dimer Stability—Once formed, the initial SL1–SL2 dimer can undergo subsequent conformational changes via the following two unimolecular rearrangements: formation of either the PAL1 or PAL2 intermolecular duplexes. Overall K_{dimer} values for formation of the PAL1, PAL2, or final dimer are taken to be the dimerization affinities for the PAL2Del, PAL1Stb, and native sequence dimers, respectively. We can then calculate the incremental, or stepwise, contribution of forming the PAL1 or PAL2 intermolecular duplexes relative to the SL1–SL2 dimer using the following relationship: overall K_{dimer} = (product of K_x values for pathway).

Formation of the PAL1 or PAL2 intermolecular duplex from the SL1–SL2 dimer state is characterized by similar incremental dissociation constants of 0.021 and 0.018, respectively (K_2 and K_2'; Fig. 8). These microscopic constants correspond to stabilizing free energy increments of −2 kcal/mol.

The net thermodynamic increment for forming the PAL1 or PAL2 intermolecular duplex, as an isolated structural element, can be calculated as the difference in free energy of the monomer and dimer states (Scheme 1). The PAL1 hairpin in the monomer and its intermolecular duplex forms have calculated (35) free energies of −4.6 and −14.1 kcal/mol; dimerization at PAL1 therefore could, in principle, contribute −9.5 kcal/mol to the stability of the dimer. PAL2 lies within a larger flexible domain that is well modeled as three interconverting constituent structures (15). The net increment for forming the PAL2 duplex in the dimer from these three structures is approximately −9.3 kcal/mol.

The calculated contributions for formation of either the PAL1 or PAL2 intermolecular duplexes as isolated secondary structure elements relative to their structures in the monomer state are therefore each greater than −9 kcal/mol based on nearest neighbor free energy parameters (35, 39) (Scheme 1). There exists a large discrepancy between these calculated values and the measured increment (Fig. 3) of only −2 kcal/mol.

Once either the PAL1 or PAL2 intermolecular duplex forms initially, formation of the second intermolecular duplex makes an even more modest contribution to dimer stability. K_3 and K_3' are 0.49 and 0.42 for PAL1 and PAL2 intermolecular duplex formation, respectively (Fig. 8). The corresponding free energy increment is −0.5 kcal/mol. Again, this increment is much smaller than that expected for formation of isolated PAL1 and PAL2 duplexes.

In the context of the retroviral MiDAS RNA, formation of the extended intermolecular duplexes at PAL1 and PAL2 in the dimer state must nearly balance disruption of favorable interactions in the monomer state. The net energetic increment for forming these new base pairing interactions is remarkably small, roughly equal to that of a single base pair (39).

PAL2 and the SL1–SL2 Domain Interactions Are Structurally Coupled—Two of the major dimerization steps involve independent interactions between autonomous secondary structural elements. In the first step, the SHAPE reactivity profile for the Stb-Del mutant supports a model in which the structure of the initial SL1–SL2 dimer strongly resembles that of the starting monomer state (compare Figs. 1*A* and 6*C*). Moreover, the dimerization dissociation constant for a truncated RNA in the monomer-like state spanning just the SL1–SL2 domain is 11 nM (Fig. 7*A*) (16) or within 2.3-fold of our measurements of the SL1–SL2 dimer in the context of the intact MiDAS RNA (corresponding to the Stb-Del mutant). Similar binding constants provide independent support for the model that SL1–SL2 domain interactions are stable and form independently of other interactions in the dimer.

Similarly, SHAPE analysis of the PAL1 dimer intermediate (illustrated by the PAL2Del mutant) indicates that PAL1 forms without affecting other elements of the structure. Specifically, the anchoring helix does not melt; PAL2 still lies in the flexible domain, and SL1 remains in a monomer-like conformation (Fig. 8, *PAL1 dimer*). This suggests that PAL1 duplex formation does not involve significant rearrangements of other parts of the MiDAS RNA.

In contrast, SHAPE analysis of the PAL2 dimer (quantified as the PAL1Stb mutant) indicates that PAL2 intermolecular duplex formation is coupled to two other processes: melting of the anchoring helix and elongation of SL1 by 4 bp to form a structure similar to that found in the final dimer state (Fig. 8, *PAL2 dimer*). PAL2 intermolecular duplex formation is therefore not structurally autonomous, as is PAL1 dimerization. Extension of SL1 by 4 bp yields a distinctive SL1–SL2 domain dimer complex whose affinity has been measured independently to be 0.3 nM (Fig. 7*B*) (16).

Thus, the favorable change in the dimerization dissociation constant for formation of the PAL2 dimer also reflects the contribution of the conformational change in the SL1–SL2 domain (Fig. 8, *black box*). Dimerization dissociation constants for the SL1–SL2 domain (in the final dimer conformation; Fig. 7*B*) and for the final, intact MiDAS dimer (Fig. 8) are identical, within error. The energetic contribution for base pairing in PAL2 is likely even smaller than the observed −2 kcal/mol increment. Instead, formation of PAL2 is linked to extension of SL1 and formation of a high affinity state in the SL1–SL2 domain.

Implications for Retroviral Biology—Extensive prior work has emphasized that the ability to form extended intermolecular duplexes, like PAL1 and PAL2, is a conserved feature of retroviral genomic dimerization domains (5, 6, 19, 22, 40). Similarly, stable stem-loop structures containing terminal GACG tetraloops are structurally conserved features of the gamma (19) and nonprimate lenti (29) retroviruses. These structures are important for RNA encapsidation (25), viral infectivity (28), and dimer formation (3) *in vivo*. Heterologous RNAs carrying just the SL1–SL2 domain are sufficient to enable some packaging of retroviral RNAs into virions (3, 28). Structurally, gamma retroviral RNAs carrying the SL1–SL2 domain form weak, but detectable, "pre-immature" dimers (3).

However, RNAs spanning only the SL1–SL2 region do not encompass the sequences adequate for highly stringent dimerization *in vitro* (15) or for optimal infectivity (3, 18) and Gag protein binding (41) *in vivo*. We reconcile this diverse information by proposing that the SL1–SL2 domain is the primary determinant for dimer stability. The highly conserved ability of retroviral dimerization domains to form intermolecular duplexes then functions to facilitate other processes of the replication cycle such as enhancing the packaging of retroviral genomic RNA into nascent virions, modulating recognition by the viral Gag or nucleocapsid proteins, or regulating viral maturation.

Acknowledgment—We are indebted to Andrew Kaplan for many insightful discussions.

REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
2. Fu, W., and Rein, A. (1993) *J. Virol.* **67,** 5443–5449
3. Hibbert, C. S., Mirro, J., and Rein, A. (2004) *J. Virol.* **78,** 10927–10938
4. Sakuragi, J., Shioda, T., and Panganiban, A. T. (2001) *J. Virol.* **75,** 2557–2565
5. Paillart, J. C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004) *Nat. Rev. Microbiol.* **2,** 461–472
6. D'Souza, V., and Summers, M. F. (2005) *Nat. Rev. Microbiol.* **3,** 643–655
7. Hu, W. S., and Temin, H. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 1556–1560
8. Stuhlmann, H., and Berg, P. (1992) *J. Virol.* **66,** 2378–2388
9. Mikkelsen, J. G., Rasmussen, S. V., and Pedersen, F. S. (2004) *Nucleic Acids Res.* **32,** 102–114
10. Rasmussen, S. V., and Pedersen, F. S. (2004) *Virology* **329,** 440–453
11. Bender, W., and Davidson, N. (1976) *Cell* **7,** 595–607
12. Bender, W., Chien, Y. H., Chattopadhyay, S., Vogt, P. K., Gardner, M. B., and Davidson, N. (1978) *J. Virol.* **25,** 888–896
13. Murti, K. G., Bondurant, M., and Tereba, A. (1981) *J. Virol.* **37,** 411–419
14. Hibbert, C. S., and Rein, A. (2005) *J. Virol.* **79,** 8142–8148
15. Badorrek, C. S., and Weeks, K. M. (2005) *Nat. Chem. Biol.* **1,** 104–111
16. Badorrek, C. S., Gherghe, C. M., and Weeks, K. M. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 13640–13645
17. Badorrek, C. S., and Weeks, K. M. (2006) *Biochemistry* **45,** 12664–12672
18. Adam, M. A., and Miller, A. D. (1988) *J. Virol.* **62,** 3802–3806
19. Konings, D. A., Nash, M. A., Maizel, J. V., and Arlinghaus, R. B. (1992) *J. Virol.* **66,** 632–640
20. Tounekti, N., Mougel, M., Roy, C., Marquet, R., Darlix, J., Paoletti, J., Ehresmann, B., and Ehresmann, C. (1992) *J. Mol. Biol.* **223,** 205–220
21. Girard, P. M., Bonnet-Mathoniere, B., Muriaux, D., and Paoletti, J. (1995) *Biochemistry* **34,** 9785–9794
22. Paillart, J., Marquet, R., Skripkin, E., Ehresmann, C., and Ehresmann, B. (1996) *Biochimie* (*Paris*) **78,** 639–653
23. Oroudjev, E. M., Kang, P. C. E., and Kohlstaedt, L. A. (1999) *J. Mol. Biol.* **291,** 603–613
24. Ly, H., and Parslow, T. G. (2002) *J. Virol.* **76,** 3135–3144
25. Yang, S., and Temin, H. M. (1994) *EMBO J.* **13,** 713–726
26. Kim, C., and Tinoco, I. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9396–9401
27. Rasmussen, S. V., Mikkelsen, J. G., and Pedersen, F. S. (2002) *J. Virol.* **323,** 613–628
28. Mougel, M., and Barklis, E. (1997) *J. Virol.* **71,** 8061–8065
29. Monie, T. P., Greatorex, J. S., Maynard-Smith, L., Hook, B. D., Bishop, N., Beales, L. P., and Lever, A. M. (2005) *Biochemistry* **44,** 294–302
30. Ly, H., Nierlich, D., Olsen, J., and Kaplan, A. (1999) *J. Virol.* **73,** 7255–7261
31. Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) *J. Am. Chem. Soc.* **127,** 4223–4231
32. Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2005) *J. Am. Chem. Soc.* **127,** 4659–4667
33. Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2006) *Nat. Protocols* **1,** in press
34. Das, R., Laederach, A., Pearlman, S. M., Herschlag, D., and Altman, R. B. (2005) *RNA* (*N. Y.*) **11,** 344–354
35. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 7287–7292
36. De Tapia, M., Metzler, V., Mougel, M., Ehresmann, B., and Ehresmann, C. (1998) *Biochemistry* **37,** 6077–6085
37. D'Souza, V., Dey, A., Habib, D., and Summers, M. F. (2004) *J. Mol. Biol.* **337,** 427–442
38. D'Souza, V., and Summers, M. F. (2004) *Nature* **431,** 586–590
39. Turner, D. H., and Sugimoto, N. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 167–192
40. Berkowitz, R., Fisher, J., and Goff, S. P. (1996) *Curr. Top. Microbiol. Immunol.* **214,** 177–218
41. Evans, M. J., Bacharach, E., and Goff, S. P. (2004) *J. Virol.* **78,** 7677–7684