2014

Structure and dynamics of the HIV-1 frameshift element RNA

Kathryn Mouzakis
Loyola Marymount University, kathryn.mouzakis@lmu.edu

Recommended Citation

Low, Justin T et al. "Structure and dynamics of the HIV-1 frameshift element RNA." Biochemistry vol. 53,26 (2014): 4282-91. doi:10.1021/bi5004926

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Structure and Dynamics of the HIV-1 Frameshift Element RNA

Justin T. Low,[†,‡] Pablo Garcia-Miranda,[§] Kathryn D. Mouzakis,[§] Robert J. Gorelick,[∥] Samuel E. Butcher,[§] and Kevin M. Weeks*[,†]

[†]Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, United States

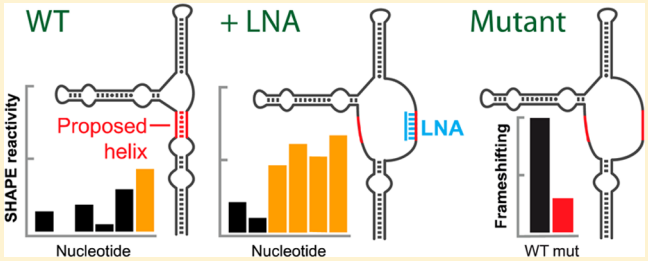
[‡]Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina 27599-7260, United States

[§]Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706-1544, United States

[∥]AIDS and Cancer Virus Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland 21702-1201, United States

Supporting Information

ABSTRACT: The HIV-1 ribosomal frameshift element is highly structured, regulates translation of all virally encoded enzymes, and is a promising therapeutic target. The prior model for this motif contains two helices separated by a three-nucleotide bulge. Modifications to this model were suggested by SHAPE chemical probing of an entire HIV-1 RNA genome. Novel features of the SHAPE-directed model include alternate helical conformations and a larger, more complex structure. These structural elements also support the presence of a secondary frameshift site within the frameshift domain. Here, we use oligonucleotide-directed structure perturbation, probing in the presence of formamide, and in-virion experiments to examine these models. Our data support a model in which the frameshift domain is anchored by a stable helix outside the conventional domain. Less stable helices within the domain can switch from the SHAPE-predicted to the two-helix conformation. Translational frameshifting assays with frameshift domain mutants support a functional role for the interactions predicted by and specific to the SHAPE-directed model. These results reveal that the HIV-1 frameshift domain is a complex, dynamic structure and underscore the importance of analyzing folding in the context of full-length RNAs.



HIV-1 viral enzymes, including protease, reverse transcriptase, and integrase, are generated by cleavage of the precursor polyprotein Gag-Pol. Although the Gag and Pol polyproteins are translated together, the *pol* gene is encoded in a reading frame offset from the upstream *gag* reading frame by one nucleotide in the 5′ direction. Consequently, translation of the Gag-Pol fusion protein relies on programmed ribosomal frameshifting that involves translation of the Gag polyprotein followed by a recoding event that shifts the ribosome from the *gag* to the *pol* reading frame.[1] Frameshifting requires two key elements: The first is a highly conserved UUUUUUA sequence, termed the slippery sequence, at which the switch in reading frame occurs,[2] and the second is a downstream structural element termed the frameshift stimulatory stem. This downstream structure is thought to pause the ribosome while the A and P sites are occupied by the slippery sequence. Although the precise mechanism is not fully understood, frameshifting occurs with a frequency of 5−10% in cultured HIV-transfected human cells, and the Gag to Gag-Pol ratio appears to be important for viral fitness.[3] The frameshifting process has consequently attracted interest as a target for the development of therapeutic agents.[4,5]

The original model for the downstream structural element was a single stem-loop.[6] A number of refinements and extensions of this model have been proposed that include additional sequence and structures including pseudoknots,[7−9] a triple-stranded RNA species,[10] and two helices.[11] NMR studies performed on 41- and 45-nucleotide (nt) transcripts support the formation of the two-helix model.[12,13] This conventional two-helix model includes the originally proposed stem (now termed the upper stem) and a lower stem that is separated from the upper stem by a three-purine bulge (Figure 1). The functional importance of this lower stem is supported by experiments demonstrating decreased frameshifting when the lower stem is destabilized by mutation[11] or when a truncated construct containing only the upper stem is used.[14]

An alternative, more complex model was proposed based on SHAPE chemical probing experiments performed on an entire HIV-1 genome.[15] SHAPE probing yields a model-free measurement of local nucleotide flexibility that, in turn, provides nucleotide-resolution information about RNA secondary and tertiary structure.[16] SHAPE reactivities can be incorporated into thermodynamics-based folding algorithms[17] resulting in highly accurate RNA secondary structure models.[18,19] The SHAPE-directed model of the frameshift domain includes five helices (Figure 1). One is equivalent to the upper stem of the

Received: April 24, 2014
Revised: June 11, 2014
Published: June 13, 2014

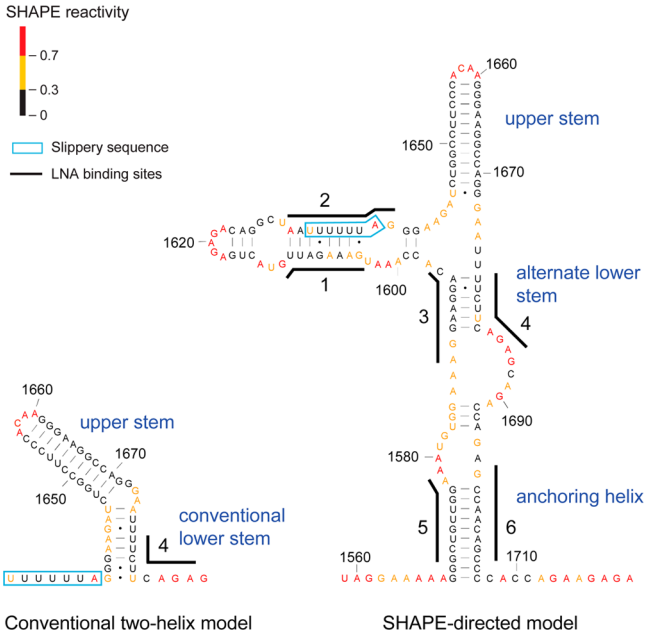


Figure 1. Conventional[11] and SHAPE-directed[15] models of the frameshift RNA element. Nucleotides are colored by SHAPE reactivity (see key) of the *ex vivo* RNA and are numbered relative to the NL4-3 genome.

two-helix model. One of the additional helices involves refolding of the lower stem of the conventional model into a helix we term the alternate lower stem. SHAPE data also support the formation of three helices outside the domain traditionally identified as the frameshift stimulatory element. These include a helix that sequesters most of the slippery sequence in base-pairing interactions and a 10-bp anchoring helix. Finally, low SHAPE reactivity at a 3-nt strand between the alternate lower stem and anchoring helix supports an additional short secondary structure, which we currently model as a 3-bp helix. SHAPE data suggest that the frameshift domain spans 140 nts, a significantly larger region than is included in the two-helix model (Figure 1).

We used several strategies to explore the relative roles of helices in the SHAPE-directed and two-helix models. We used oligonucleotides containing locked nucleic acid (LNA) residues[20,21] to selectively disrupt predicted helices. We also used SHAPE to monitor frameshift element structure in the presence of formamide denaturant and in the context of the RNA packaged inside virion particles. To assess the functional importance of the frameshift domain helices, we used site-directed mutagenesis to destabilize individual helices and measured the resulting frameshift efficiency.[22] Our results provide strong support for the SHAPE-directed frameshift model, and the functional importance of SHAPE-detected conformations for frameshifting reveals that the frameshift domain is a dynamic element capable of structural remodeling and supports the existence of a secondary frameshift site. These results were critically dependent on the use of the entire HIV-1 sequence, a fact that emphasizes the importance of sequence context for studying large, complex functional domains in RNA.

METHODS

HIV-1 Virion Production. HIV-1 virion particles were prepared as described.[23] Briefly, HIV-1 NL4-3 (group M, subtype B) derived by the transfection of 293T cells with pNL4-3 (obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3 from Dr. Malcolm Martin[24]) was used to infect cells from a non-Hodgkin's T cell lymphoma line (SupT1-CCR5).[25] Purified HIV-1 preparations (1000× concentrates) were stored frozen at −80 °C until used.

***In Virio* RNA Modification and Extraction of RNA Genomes from Virions.** Virions were treated as described,[26] with the exception that the 1M7 SHAPE reagent was used in place of NMIA. Briefly, virions treated with 1M7 or unmodified virions were digested with subtilisin and centrifuged through a 20% (w/v) sucrose cushion prior to the extraction of genomes.[27] Virions were then lysed by incubation in virion lysis buffer [50 mM Tris (pH 7.5), 10 mM EDTA, 1% SDS, 100 mM NaCl, 10 mM DTT, 15 μL 20 mg/mL proteinase K] (using 100 μL lysing buffer per 1 mL volume of 1000× HIV-1 culture supernatant concentrate) for 30 min at room temperature. The digest was extracted four times with phenol/chloroform/isoamyl alcohol, followed by four extractions with pure chloroform. The aqueous layer was brought to a NaCl concentration of 300 mM, nucleic acids were precipitated with 70% ethanol, and samples were stored at −20 °C. Pelleted RNA was resuspended in storage buffer (50 mM HEPES, pH 8) at a final RNA concentration of 400 nM. These aliquots were flash frozen in liquid N_2 and stored at −80 °C.

Folding of *Ex Virio* RNA. For each reaction, 1 μL of 400 nM *ex vivo* RNA (prepared as described above, without *in vivo* RNA chemical modification) was dissolved in a Mg^{2+}-containing standard folding buffer [50 mM HEPES (pH 8), 200 mM potassium acetate (pH 8), and 3 mM $MgCl_2$] to promote the formation of native-like interactions (total volume was 20 μL). This mixture was incubated at 37 °C for 30 min before SHAPE modification.

Formamide Denaturation. For the formamide denaturation experiments, 1 μL of 400 nM *ex vivo* RNA was resuspended in 5 μL of folding buffer and incubated at 37 °C for 20 min. A formamide-containing folding buffer [67% (v/v) deionized formamide, 50 mM HEPES (pH 8), 200 mM potassium acetate (pH 8), and 3 mM $MgCl_2$] was then combined with the appropriate amount of standard folding buffer and added to the reaction mix to obtain the desired formamide concentration in a reaction volume of 20 μL. Experiments were performed at final formamide concentrations of 0% to 60%, in increments of 10%. The RNA was incubated in formamide for 20 min at 37 °C prior to SHAPE modification.

LNA Oligonucleotide Design. We used 9- and 10-nt LNA oligonucleotides (Exiqon) that were the reverse complements of HIV-1 sequences comprising helices proposed in the SHAPE-directed frameshift model. The sequences were chosen to avoid self-complementarity and to avoid stretches of three or more G or C nucleotides. When possible, we included flanking single-stranded regions as part of the targeted strand to facilitate LNA binding to structured helices. LNA oligonucleotide binding positions (in parentheses) refer to the NL4-3 genome (based on GenBank accession number AF324493, starting at the beginning of the 5′ R-region, nucleotide 455), and the sequences are 1 (1604−1612), CAATCTTTC; 2 (1629−1638), CTAAAAAATT; 3 (1588−1596), GTCCTTCCT; 4 (1678−1686), TCTGAAGAA; 5 (1570−1578), TCCAACAGC; and 6 (1697−1705), GCTGTTGGC. LNA nucleotides are underlined; all others are DNA. Their corresponding binding sites are shown in Figure 1.

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

LNA Binding to Genomic RNA. For each LNA binding experiment, 1 μL of 400 nM extracted HIV-1 genomic RNA in storage buffer [50 mM HEPES (pH 8) and 200 mM potassium acetate], 2 μL of 2 μM LNA, 4.6 μL of TE, and 1.7 μL of water were mixed and heated at 95 °C for 5 min and then snap cooled on ice. A 9.5 μL aliquot of 2× storage buffer and 1.2 μL of 50 mM $MgCl_2$ (20 μL total in folding buffer conditions) were added and the solution incubated for 30 min at 37 °C before the SHAPE modification step.

SHAPE Modification of *Ex Virio*, LNA-Bound, and Formamide-Denatured RNA. SHAPE modification was performed by adding 9 μL each of RNA sample to 1 μL of 40 mM SHAPE reagent (1M7 for standard SHAPE reactions and 1M6 and NMIA for differential SHAPE reactions) or 1 μL of neat DMSO as a no-reagent control. The samples were incubated for 4 min at 37 °C. We then added 1 μL of 50 mM EDTA and performed a cleanup step to remove LNA oligonucleotides if necessary (RNeasy Mini Kit; Qiagen). The resulting mixture (100 μL) was precipitated with 10 μL of 2 M NaCl and 300 μL of ethanol and stored at −80 °C until primer extension.

Primer Extension and Capillary Electrophoresis Detection of SHAPE Adduct Sites. Pellets containing SHAPE-modified or control RNA were resuspended in 7 μL of 0.5× TE and 6 μL of 0.4 μM VIC-labeled DNA primer (Applied Biosciences). Primer extension reactions for LNA binding experiments were performed using DNA primers complementary to RNA genome nucleotide positions (see above) 1750−1771 (primer 6.2, with sequence ATC GGC TCC TGC TTC TGA GAG G) and 2033−2054 (primer 7, with sequence CAA TTA TGT TGA CAG GTG TAG G). We used primer 6.2 for formamide denaturation experiments and primer 7 for *in virio* experiments. Primer extension reactions were performed as described.[28,29]

Data Processing. Raw capillary electropherograms were processed using QuShape.[28] Briefly, key processing steps include a mobility shift to correct for small differences in the electrophoretic mobility between the NED and VIC fluorescent dyes and a signal decay correction to account for signal attenuation as the distance from the reverse transcriptase primer binding site increases. DMSO control peaks were subtracted from reagent peaks, and the resulting SHAPE reactivities were normalized on a scale where a normalized reactivity of 1.0 was defined as the average intensity of the top 10% most reactive peaks, excluding a few highly reactive nucleotides taken to be outliers. The resulting reactivities span a scale from 0 to ∼1.5, where 0 indicates no reactivity (and a highly constrained nucleotide) and reactivities >0.7 typically indicate highly flexible nucleotides.

Secondary Structure Modeling. SHAPE-directed models for the frameshift domain were created by incorporating SHAPE data into the RNAstructure folding algorithm.[17,18] For HIV, we used values of m = 3.0 kcal/mol and b = −0.6 kcal/mol. To account for LNA binding, the LNA target site was forced to be single stranded by imposing artificial SHAPE reactivity values of 100. For LNA 5 and 6, which target the highly stable anchoring helix, base pairing was also prohibited at partner nucleotides as the observed reactivity increases were taken to imply a single stranded state.

Plasmid Construction. Complementary synthetic oligonucleotides (IDT) with BamH I and Sac I compatible ends were phosphorylated, annealed, and ligated into the p2luc vector using the BamH I and Sac I sites between the rluc and fluc reporter genes.[22] Plasmid DNA was purified from cell cultures (Qiagen), and the sequences of all constructs were verified by sequencing.

RNA Synthesis, Purification, and the Frameshift Assay. RNAs for the frameshift assay were transcribed *in vitro* from p2luc plasmid DNA linearized with PmlI, purified His6-tagged T7 RNA polymerase (10×), 11.25 mM NTPs, and 2 units of RNase inhibitor (RNasin Plus; Promega, N2615) in 200 μL for 2 h at 37 °C. Pyrophosphate was pelleted by centrifugation (10 min, 13,200 rpm), and RNA was extracted with phenol/chloroform. Unincorporated NTPs and salt were separated from the RNA by gel filtration (Ilustra MicroSpin G-25 Columns; GE Healthcare). RNAs were heated at 95 °C for 5 min followed by incubation on ice for 30 min. Samples were lyophilized to dryness and resuspended in water at 1 μg/μL. RNA integrity and purity were confirmed with 1% agarose gel electrophoresis.

In vitro frameshift assays were completed with each RNA reporter using a rabbit reticulocyte lysate (RRL System; Promega, nuclease treated, L416A). Translation reactions contained 0.63 μg of RNA, 10 units of RNase inhibitor, and 8.8 μL of RRL in 12.5 μL. Following a 90 min incubation at 37 °C, reactions were quenched by the addition of EDTA (pH 8.0) to a final concentration of 6 mM. For each reporter, a minimum of three independent frameshift assays were completed using different preparations of mRNA (full biological replicates). Independent biological replicate assays included four replicate reactions. Luminescence was measured using the dual-luciferase reporter assay (Promega). Readings were taken with a microplate luminometer (Veritas) equipped with dual-injectors (Turner Biosystems) for 10 s after 25 μL of the respective substrate was injected into the reaction mixture (2 s lag time prior to measurement). Each experiment included an in-frame positive control[22] with a mutated "non-slippery sequence" and an additional nucleotide inserted immediately before the Sac I site, which places the rluc and fluc genes in-frame. Frameshift efficiencies were calculated by taking the ratio of the experimental to control luminescence (firefly to *Renilla*). Small variations in the activities of control reactions were observed with different preparations of mRNA and reticulocyte lysates. To account for these differences, frameshift efficiencies were normalized relative to wild-type levels, and averaged. The non-normalized frameshift efficiencies for the native sequence construct in all biological replicates (n = 9) was 5.4 ± 0.9%. Standard deviations were propagated to yield a standard error of the mean (SEM).

RESULTS

LNA Binding Supports the SHAPE-Directed Model of the Frameshift Region. We initially attempted to bind antisense LNAs to folded full-length genomic RNA extracted from HIV-1 virions. However, addition of LNAs directed against the anchoring helix failed to produce detectable SHAPE reactivity changes in the predicted partner strand. This result could reflect either that the targeted helix was too stable to be disrupted by LNA binding or that our model in the frameshift region was inaccurate. When we heated the HIV-1 RNA at 95 °C for 5 min in the presence of LNA and snap cooled on ice prior to incubation in folding buffer, we found that SHAPE reactivities did increase in the predicted partner strand upon addition of LNA, indicative of a very stable target helix.

For the results obtained upon heat denaturation and LNA binding in the frameshift element region to provide clear

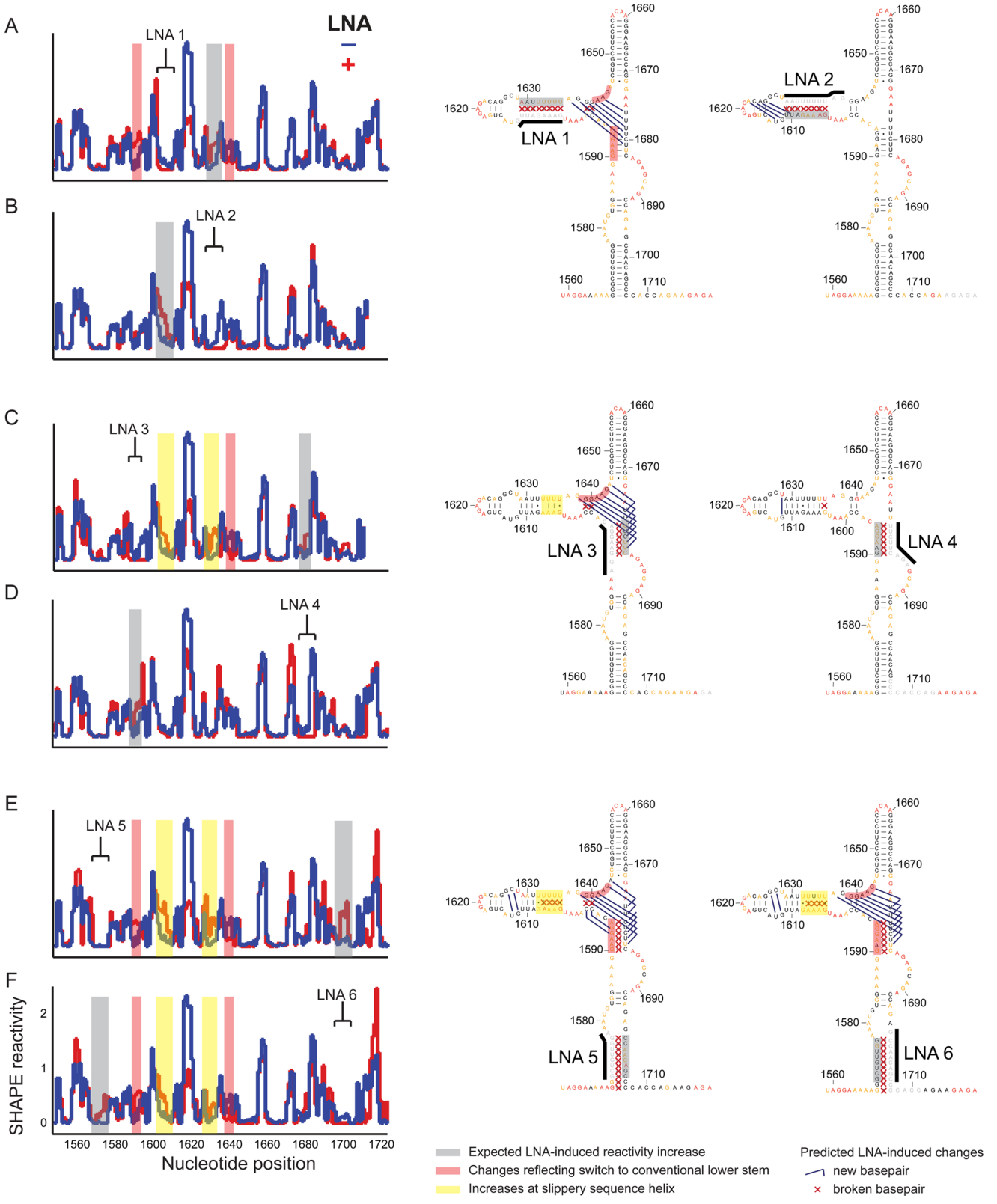


Figure 2. SHAPE reactivities and predicted secondary structures for HIV-1 RNA bound to LNAs targeting the slippery sequence helix (A,B), the SHAPE-supported alternate lower stem (C,D), and the anchoring helix (E,F).

biological insights, the conformation adopted after heat denaturation must be native-like. We performed SHAPE on HIV-1 RNA gently extracted from virions and on a sample of this same RNA following heat denaturation and refolding. The standard SHAPE reagent 1M7[30] and two additional SHAPE probing reagents, 1M6 and NMIA, were used in these analyses. 1M6 and NMIA are sensitive to different local structural states and time scales and consequently provide a detailed structural fingerprint for an RNA.[31,32] SHAPE reactivities for the genome region near the frameshift element were highly similar in both

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

ex virio and in heated-refolded RNAs for all three reagents (r = 0.94−0.96, Supporting Information, Figure 1). We conclude that the refolded frameshift domain adopts essentially the same conformation as the *ex virio* HIV-1 RNA purified under nondenaturing conditions.

We designed 9- and 10-nt LNA oligonucleotides to bind to each of the strands of the three major helices unique to the SHAPE-directed frameshift model (Figure 1). If the targeted helix existed and if LNA binding disrupted native base pairing, we expected to observe increased SHAPE reactivity at the strand complementary to the LNA-bound strand. In all six cases, SHAPE reactivity increased for at least some nucleotides in the predicted partner strand (Figure 2, gray boxes). These results strongly support the formation of the slippery sequence (Figure 2A,B) and alternate lower (Figure 2C,D) and anchoring helices (Figure 2E,F) in the full-length HIV-1 RNA.

LNA Binding Allows the Conventional Lower Stem to Form. Although reactivity increases were consistently observed at the predicted partner strand upon LNA binding, most LNAs also induced additional SHAPE reactivity changes. For example, LNA 1 was designed to disrupt the slippery sequence helix. In addition to the expected increases in SHAPE reactivity at the partner strand (Figure 2A, gray boxes), LNA 1 binding also induced SHAPE reactivity increases at nucleotides 1585−1590 and decreases at nucleotides 1641−1644. Nucleotides 1585−1590 are predicted to be single stranded in the SHAPE-directed frameshift model, while nucleotides 1641−1644 are predicted to form part of the alternate lower stem. The changes in SHAPE reactivity upon LNA binding suggest the formation of base-pairing interactions that correspond to the lower stem from the conventional two-helix model (Figure 2A, red boxes). Binding to the complementary side of the helix by LNA 2 did not result in this conformational switch (Figure 2B). LNA 2 targets nucleotides in close proximity to, and possibly overlapping with, the conventional lower stem. In this case, steric occlusion due to LNA 2 binding likely disfavors formation of the conventional lower stem.

LNA 3 was designed to disrupt the alternate lower stem. LNA 3 induced relatively minor increases in reactivity in the partner strand (Figure 2C, gray box) and more pronounced changes in reactivity in other regions (Figure 2C, red and yellow boxes). For example, reactivity at nucleotides 1641−1643 was reduced relative to the reactivity in the no-LNA control, consistent with the formation of the conventional lower stem. LNA 4 was also designed to disrupt the alternate lower stem, and its binding resulted in the expected reactivity increases at the predicted partner strand but no significant changes elsewhere. We attribute this lack of a conformational switch to the fact that LNA 4 binding would prevent the formation of both the alternate and conventional conformations of the lower stem.

Binding by LNA 5 or LNA 6 also resulted in SHAPE reactivity increases in the corresponding partner strands (Figure 2E,F) and in changes similar to those observed upon LNA 1 binding to the slippery sequence helix and LNA 3 binding to the alternate lower stem. SHAPE reactivities decreased for nucleotides 1641−1643 and increased for nucleotides 1590−1595, consistent with a transition from the alternate lower stem to the conventional lower stem. These results suggest that the anchoring helix stabilizes the alternate lower stem.

Formation of the Conventional Lower Stem Destabilizes the Slippery Sequence Helix. Binding by LNAs 3, 5, and 6 all induced a switch in RNA conformation from the alternate lower stem to the conventional lower stem (Figure 2C,E,F, red boxes). Binding by these LNAs also resulted in increased SHAPE reactivity at the slippery sequence helix (Figure 2C,E,F, yellow boxes). Remarkably, these changes even occurred upon disruption of the relatively distant anchoring helix by LNAs 5 and 6. In contrast, LNA 4, which destabilized the alternate lower stem but did not induce the formation of the conventional lower stem, did not cause significant SHAPE reactivity changes in the slippery sequence helix (Figure 2D). These results suggest that the alternate lower stem and slippery sequence stabilize one another and are structurally coupled.

Frameshift Helix Stability Revealed by Formamide Denaturation. We next examined the relative stabilities of the four main helices present in the SHAPE-directed frameshift element. Formamide denatures nucleic acid structure.[33] Therefore, we incubated HIV-1 RNA in buffer containing a range (0−60% vol/vol) of formamide concentrations with the expectation that global reactivity to 1M7 would increase with denaturant concentration. Following a 20 min incubation at 37 °C in folding buffer, we added formamide and incubated the solution for a further 20 min before initiating SHAPE probing with 1M7.

At 20% formamide, nucleotides involved in the slippery sequence helix showed increased SHAPE reactivity relative to the buffer without formamide (Figure 3A,B). Nucleotides involved in the alternate lower stem exhibited reactivity changes

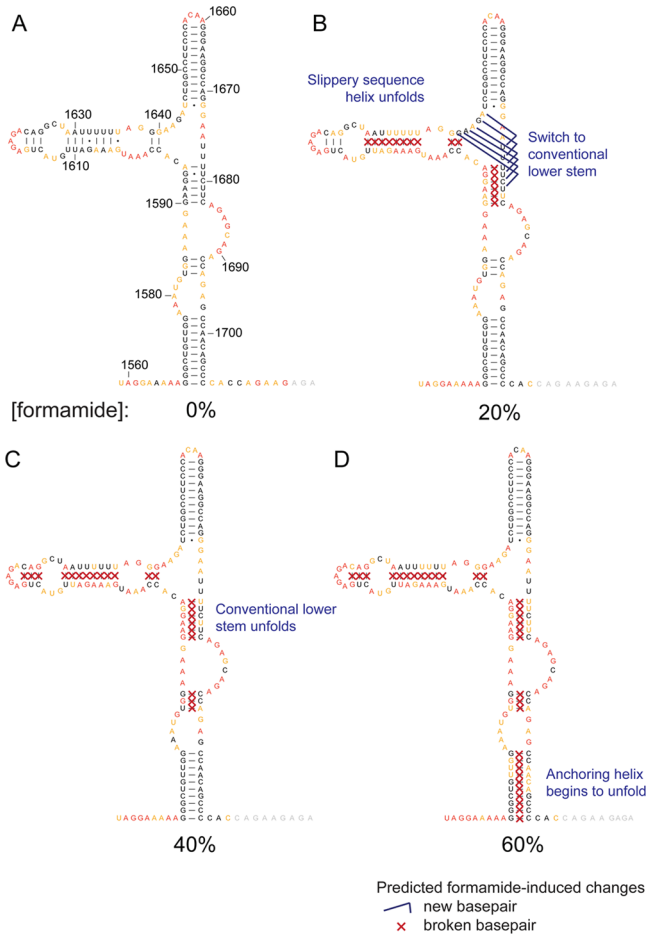


Figure 3. Secondary structure models for the frameshift domain RNA as a function of formamide concentration.

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

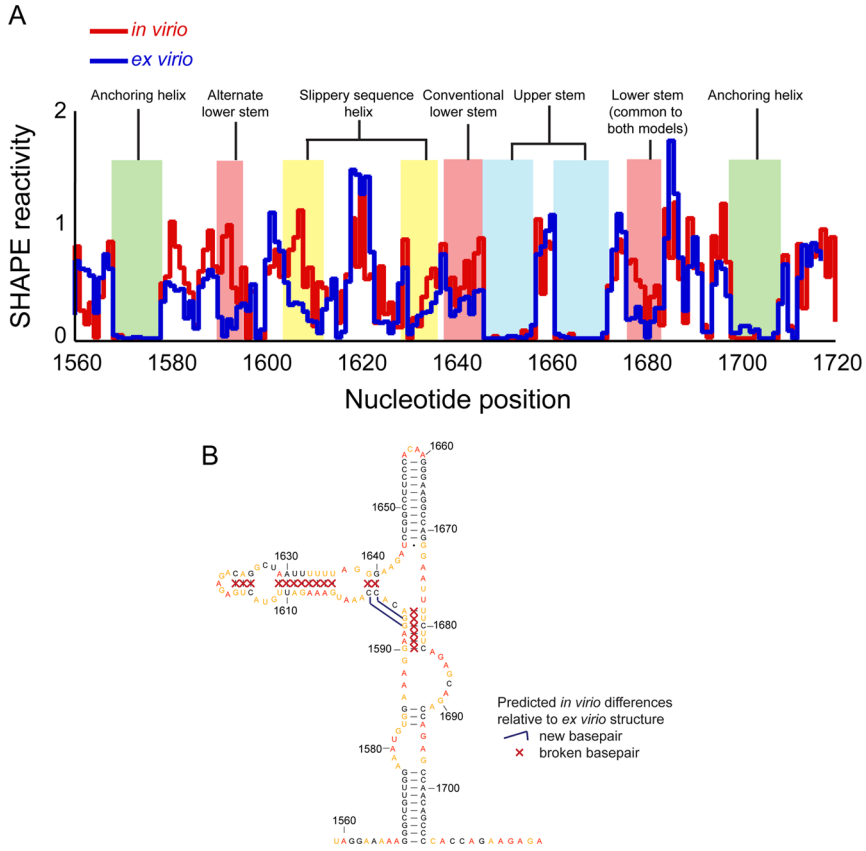


Figure 4. Structure of *in virio* compared with *ex virio* RNA. (A) SHAPE reactivity profiles and (B) predicted secondary structure model.

that suggest a switch to the conventional lower stem at 20% formamide (Figure 3B and Supporting Information, Figure 2). In 40% formamide, SHAPE reactivities indicated that the conventional lower stem was unfolded (Figure 3C). Nearest neighbor thermodynamic calculations indicate that the conventional lower stem is slightly more stable than the SHAPE-directed alternate lower stem (−8.4 kcal/mol and −8.2 kcal/mol, respectively). We hypothesize that the formation of the alternate lower stem is largely dependent on stabilization from other elements in the domain and that, when these interactions are disrupted by LNA binding or by formamide denaturation, the slightly more stable conventional lower stem is able to form.

The anchoring helix and upper stem show very low SHAPE reactivities up to 40% formamide. Nucleotides in the anchoring helix became reactive only at 60% formamide, and nucleotides in the upper stem remained unreactive even at this high denaturant concentration (Figure 3D). The high stability of the upper stem was noted previously.[34] We conclude from these experiments that the frameshift domain is bounded by two highly stable helices and that the intervening less stable helices have the propensity to unfold upon perturbation and, in the case of the lower stem, refold into alternative conformations.

Frameshift Domain Is Less Structured *In Virio* than in Isolated RNA. We next probed the frameshift region of HIV-1 RNA as it exists packaged inside virions using the well-validated 1M7 SHAPE reagent. Structural interrogation of the *in virio* state is possible because SHAPE reagents readily cross biological membranes.[35,36] SHAPE probing indicated that the frameshift domain of RNA packaged within the virion adopted a much less structured conformation than that adopted by either gently deproteinized genomic (*ex virio*) or refolded RNA (Figure 4). Nucleotides involved in the slippery sequence helix and lower stem, including both the conventional and alternate base-pairings, displayed higher SHAPE reactivities *in virio* than in isolated RNA. These differences could reflect RNA refolding or protein binding in the *in virio* environment. In contrast, SHAPE reactivities for the anchoring helix and upper stem remained low, suggesting that these elements folded stably inside virions.

SHAPE-Predicted Alternate Lower Stem Is Required for Native-Like Frameshifting. While the LNA binding experiments confirmed the existence of all helices unique to the SHAPE-directed model, the formamide denaturation and *in virio* experiments showed that some of these helices are relatively unstable. To determine the biological importance of each of these helices, we examined a set of mutants designed to disrupt individual predicted helices and evaluated frameshifting efficiency in a dual luciferase reporter assay.[22] Using this system, we were unable to detect any significant effect on frameshifting upon disruption of either the anchoring helix (Figure 5, mutants 1 and 2) or the conventional lower stem (Figure 5, mutant 3). However, when the alternate lower stem was disrupted, a significant decrease in frameshifting was observed (Figure 5, mutant 4). We conclude that native-like frameshifting levels are dependent on the formation of the alternate, but not necessarily the conventional, lower stem. These results suggest that the alternate lower stem is a potential therapeutic target.

A very different result was obtained when the anchoring helix, alternate lower stem, and slippery site helix were simultaneously disrupted by mutagenesis (Figure 5, mutant 5). In this case, frameshifting increased to greater than native-

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

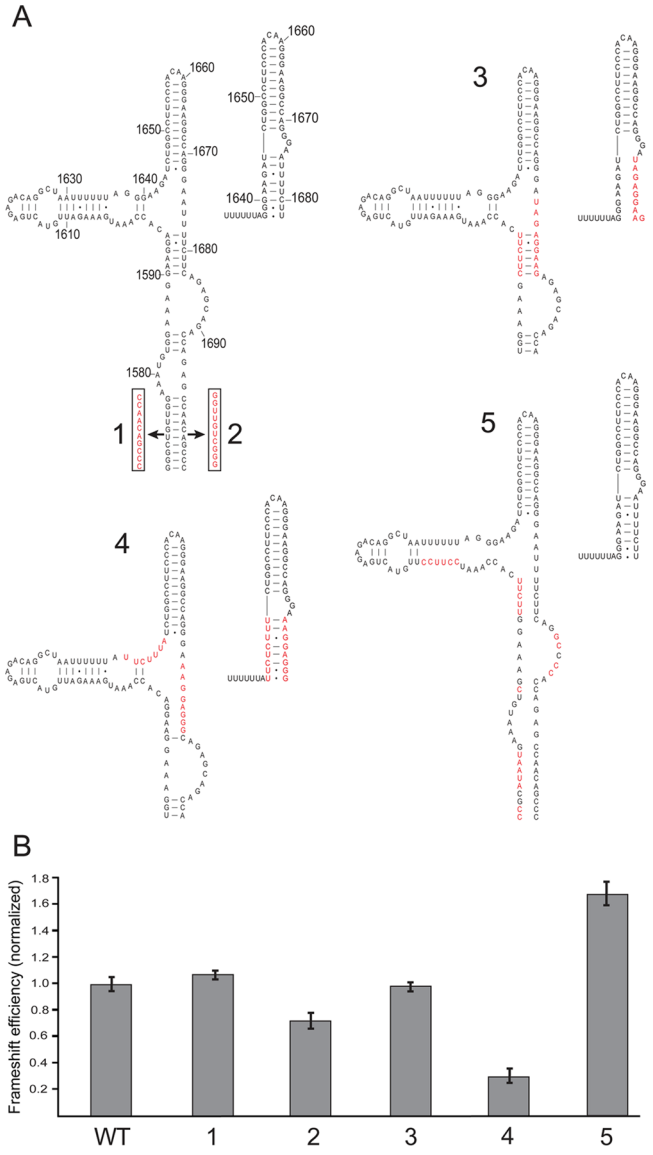


Figure 5. Frameshifting efficiencies on transcripts with mutations in predicted helical regions. (A) Native and mutant sequences and (B) relative frameshift efficiencies. Constructs were designed to disrupt the anchoring helix (1 and 2), conventional lower stem (3), SHAPE-directed alternate lower stem (4), and both the anchoring helix and alternate lower stem simultaneously (5). Mutated positions are shown in red. For mutants 3 and 4, the entire anchoring helix is present, but only 3 base pairs are shown. Frameshift efficiencies were normalized by setting the measured native sequence value of 5.4 ± 0.9% equal to 1.0.

like levels. These results agree with a prior report that frameshifting efficiency of a construct containing only the upper stem and conventional lower stem is approximately twice that of a construct that contains the entire SHAPE-predicted domain.[22] Together, these data are consistent with a translational pausing mechanism whereby structured regions preceding the frameshift site regulate ribosome spacing.[37] That these results differ from those obtained when the anchoring helix and alternate lower stem were mutated independently supports our hypotheses that there is strong structural coupling between these elements and that these helices have overlapping functional roles in modulating frameshifting efficiency.

Frameshifting at an Alternate Slippery Sequence Is Dependent on Downstream Structure. The alternate lower stem overlaps with a UUUUCUU sequence (nucleotides 1676−1682) that resembles the conserved UUUUUUA slippery sequence. In viruses resistant to protease inhibitors, the C in this downstream region is frequently mutated to U.[38] At the protein level, this causes a Leu to Phe mutation that may increase cleavage efficiency and enhance the production of functional protease. This site could also act as a secondary slippery sequence that increases the overall amount of frameshifting, thereby increasing the relative amount of protease to compensate for reduction in protease activity in the presence of inhibitor.[39] When the standard slippery sequence is inactivated by mutation, the mutant with the C to U change in the alternate slippery site, but not that containing the wild-type sequence, stimulates frameshifting in *in vitro* translation assays. In cell culture, no frameshifting was detected from the secondary site using HIV-1 sequences shorter than the 140-nt domain described here.[40] However, the construct used could not form flanking structures such as the anchoring helix and the adjacent 3-bp helix predicted by SHAPE-directed folding.

We therefore revisited frameshifting from this alternate slippery sequence in the context of the complete frameshift domain. We inactivated the standard slippery sequence by site-directed mutagenesis (Figure 6, Δss); this reduced frameshifting efficiency by 5-fold in the luciferase reporter assay. The reporter containing the C1680U mutation in the alternate slippery sequence in this Δss background (Figure 6, mutant 6) showed a 3-fold increase in frameshifting efficiency relative to the Δss construct. We then disrupted the anchoring helix (Figure 6, mutant 7) and the adjacent 3-bp helix (mutant 8) in the Δss background. Disruption of the anchoring helix had no effect on frameshift efficiency, but disruption of the 3-nt helix reduced frameshifting to levels similar to those of the Δss construct. We conclude that the C1680U mutation creates a second functional slippery sequence and that frameshifting from the secondary site requires both a slippery sequence and a downstream base-paired structure.

DISCUSSION

Here, we analyzed in detail the structure of the frameshift domain in the context of full-length HIV-1 genomic RNA. Our data provide support for the three main helices unique to the frameshift domain model proposed based on SHAPE-directed folding of an entire HIV-1 genome (Figure 1).[15] Models for the frameshift domain that contain pseudoknots have recently been proposed;[9] however, *ex virio* SHAPE data (Figure 1) and prior frameshifting assays[22] do not support a contribution of these pseudoknots to the observed structural ensemble or frameshifting efficiency. Our model shares a stable upper stem with the previously proposed two-helix model[11] but has two important differences. First, this study supports alternative pairing partners for bases in the conventional lower stem. Second, the SHAPE-directed model spans a 140-nt region that includes the slippery sequence and an additional stable anchoring helix.

A key feature of the present study is the use of full-length genomic RNA. Using a long RNA avoids the biases that are often introduced when a truncated sequence of an RNA is studied and ensures that important structural regions are not omitted. Most past studies of the HIV-1 frameshift domain, including NMR analyses that tested the two-helix model,[12,13] used HIV-1 sequences that were much shorter than the 140-nt

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

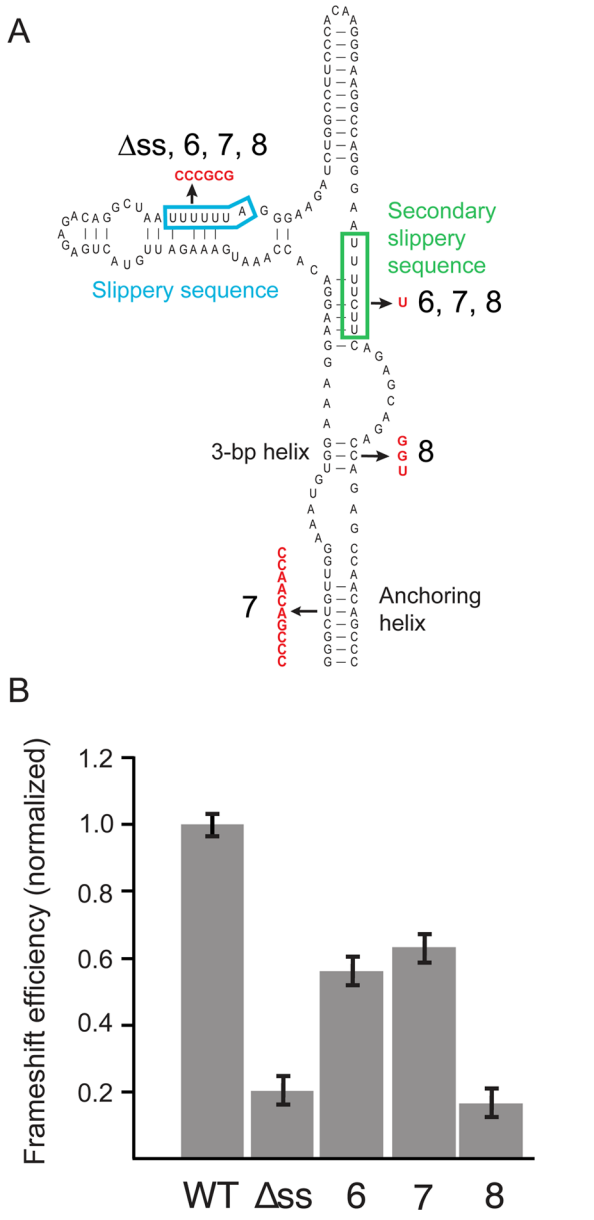


Figure 6. Frameshifting efficiencies for analysis of the secondary frameshift site. (A) Native and mutant sequences and (B) relative frameshift efficiencies. Construct Δss has mutations that disrupt the standard slippery sequence. All other constructs also contain the Δss mutation. Construct 6 additionally includes a C-to-U mutation that creates the proposed secondary slippery sequence. Constructs 7 and 8 contain this secondary slippery sequence and additional mutations that disrupt either the anchoring helix (7) or the 3-nt helix (8). Frameshift efficiencies were normalized as in Figure 5.

region that we show here, which constitutes the full frameshift domain. The three main helices unique to the SHAPE-directed model were thus unable to form in the truncated sequences analyzed previously. Two of these helices, the alternate lower stem and the helix that sequesters the slippery sequence, appear to mutually stabilize each other because disruption of either of these helices by LNA binding destabilized or induced refolding of the other (Figure 2). Additionally, the anchoring helix likely stabilizes both of these helices. The frameshift domain adopts the same structure whether the genomic RNA is probed directly after extraction from virions or is first heat-denatured and then refolded. This similarity suggests that the differences between the SHAPE-directed model and the conventional model are due primarily to the presence of the complete sequence of the viral RNA that was used in development of the SHAPE-directed model. Co-transcriptional RNA folding thus apparently does not alter the folding of this region of the genome. However, inside virions, portions of the frameshift domain are less structured than in the extracted RNA, suggesting that higher-order protein or ligand interactions modulate domain structure in the native virion environment.

The anchoring helix and upper stem are very stable and persist even in the presence of high concentrations of formamide (Figure 3). Furthermore, the *in virio* experiments show that these two helices are the only frameshift domain helices that form stably inside the virion. In the absence of tertiary structure or other potentially stabilizing interactions, the conventional lower stem is predicted to be slightly more stable than the SHAPE-directed alternate lower stem, and in 20% formamide, we observed a switch from the alternative stem to the conventional lower stem. These data suggest that the structure of the frameshift region is dependent on the local microenvironment.

The changes induced by LNA binding suggest that the structure of the frameshift domain is dynamic as the ribosome moves through this region. The anchoring helix is the first frameshift domain secondary structural element to be unwound by the translating ribosome (Figure 7A,B). LNA binding experiments indicate that unwinding of this helix induces a major structural rearrangement throughout the frameshift domain. The lower stem switches from the alternate helix to the conventional stem conformation, and the slippery sequence helix is significantly destabilized. The functional implications of this structural rearrangement are not understood, but the

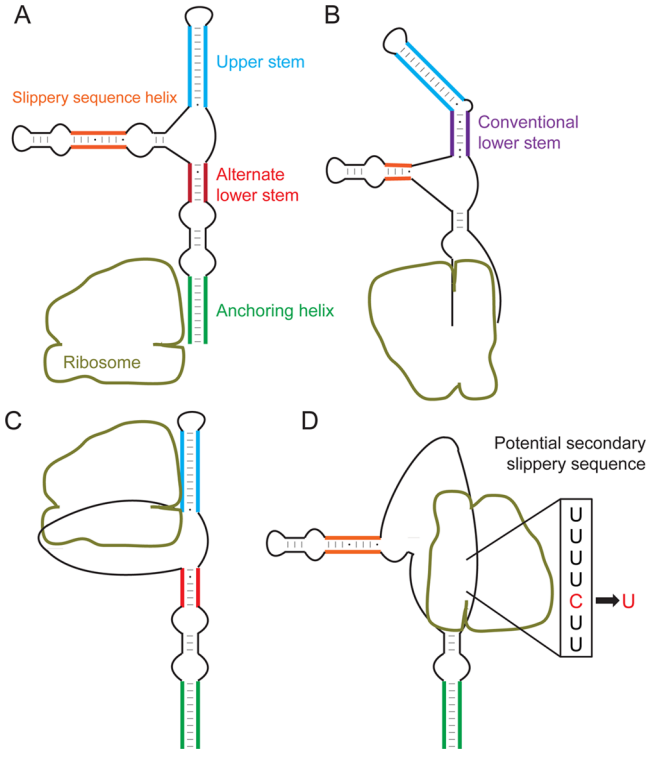


Figure 7. Model for frameshift domain unwinding during translation. A C-to-U mutation (red) observed in viruses with resistance to protease inhibitors[38] creates a potential secondary slippery sequence.

dx.doi.org/10.1021/bi5004926 | *Biochemistry* 2014, 53, 4282−4291

importance of conformational switching was recently demonstrated in ribosomal recoding in another retrovirus.[41] Prior studies using an RNA containing up to 90 nucleotides of the HIV-1 frameshift domain sequence indicate that sequences in the conventional lower stem are functionally important for frameshifting in cultured cells but not in rabbit reticulocyte lysate.[11,14] Our mutagenesis studies similarly show that disruption of the conventional lower stem does not yield detectable changes in frameshift efficiency in reticulocyte lysate (Figure 5, compare WT vs 3). Using a longer fragment of HIV-1 RNA that spans the entire SHAPE-defined domain (nts 1567−1711), we observed that disrupting the alternate lower stem while allowing the conventional lower stem to form resulted in dramatically reduced frameshifting (Figure 5, mutant 4). In contrast, when both the anchoring helix and the alternate lower stem were disrupted while preserving the ability of the conventional lower stem to form, frameshifting actually *increased* above native-like levels (Figure 5, mutant 5). Thus, the frameshift domain structure is dynamic, and the stabilities of the component helices are coupled.

After unwinding the anchoring helix, the ribosome continues toward the slippery sequence and encounters the conventional lower stem. LNA 2 binding data suggest that unwinding of the conventional lower stem results in a switch back to the alternate lower stem conformation (Figure 7C). The ribosome must then unwind the very stable upper stem and switch reading frames. The high stability of the anchoring helix suggests that this structure may reform after the ribosome has translated through it. If this is the case, the ribosome would encounter the anchoring helix for a second time before exiting the frameshift domain (Figure 7D). The adjacent 3-bp helix could also potentially reform and would be able to serve as the frameshift stimulatory stem in HIV-1 sequence variants that contain a second slippery sequence.

Frameshifting efficiency assays performed in the present study revealed that the secondary slippery site can indeed promote frameshifting. Previous studies[40] likely failed to detect frameshifting from this secondary site because of the use of truncated sequences that were unable to form downstream structures necessary for frameshifting. Frameshift efficiency is primarily dependent upon the local stability of the first three base pairs of the helix located approximately eight nucleotides from the slippery site, which is the length of RNA required to span the distance from the slippery site to the ribosomal mRNA entrance channel.[22] Consistent with this eight nucleotide distance, our results indicate that frameshifting from the second slippery site is dependent upon the integrity of the 3-bp helix but not the anchoring stem, which would be too far away from the second slippery site to induce frameshifting (Figure 6).

In sum, this study confirmed key structural elements specific to the HIV-1 frameshift region model developed using SHAPE probing data as knowledge-based restraints[18,19] to computational folding algorithms.[15] The frameshift region has significant conformational flexibility, and we hypothesize that a switch occurs between the SHAPE-predicted helices and the two helices in the conventional model as the ribosome unwinds the frameshift element. From a functional perspective, the alternate lower stem is important for native-like frameshift levels in the context of the larger domain structure and consequently represents a potential therapeutic target. We also examined the sequence and structural requirements of a previously identified secondary frameshift site and found that its function is dependent on structural elements outside the traditional frameshift domain but within the SHAPE-directed model. This work illustrates the influence of global sequence context on RNA structure and the ability of SHAPE probing to reveal complex functional domains in large RNAs.

■ ASSOCIATED CONTENT

Supporting Information

Frameshift domain SHAPE reactivities; formamide-induced destabilization of the frameshift domain and discussion. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: weeks@unc.edu.

Funding

This work was supported by National Institutes of Health grants AI068462 (to K.M.W.) and GM072447 (to S.E.B.). R.J.G was supported with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN261200800001E with Leidos Biomedical Research, Inc. J.T.L. was supported by a National Research Service Award (F30DA027364), the Medical Scientist Training Program (T32GM008719), and the Program in Molecular and Cellular Biophysics (T32GM008570).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are indebted to Howard Fried for suggesting formamide denaturation experiments and for helpful discussions, to Nathan Siegfried and Josh Martin for helpful discussions on LNA binding experiments, and to Julian W. Bess, Jr. of the AIDS and Cancer Virus Program for supplying the purified HIV-1 virions used in this study.

■ REFERENCES

(1) Brierley, I., Ramos, and Dos, F. J. (2006) Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res. 119*, 29−42.

(2) Biswas, P., Jiang, X., Pacchia, A. L., Dougherty, J. P., and Peltz, S. W. (2004) The human immunodeficiency virus type 1 ribosomal frameshifting site is an invariant sequence determinant and an important target for antiviral therapy. *J. Virol. 78*, 2082−2087.

(3) Shehu-Xhilaga, M., Crowe, S. M., and Mak, J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol. 75*, 1834−1841.

(4) Brakier-Gingras, L., Charbonneau, J., and Butcher, S. E. (2012) Targeting frameshifting in the human immunodeficiency virus. *Expert Opin. Ther. Targets 16*, 249−258.

(5) Gareiss, P. C., and Miller, B. L. (2009) Ribosomal frameshifting: an emerging drug target for HIV. *Curr. Opin Invest. Drugs 10*, 121−128.

(6) Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J., and Varmus, H. E. (1988) Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature 331*, 280−283.

(7) Le, S. Y., Shapiro, B. A., Chen, J. H., Nussinov, R., and Maizel, J. V. (1991) RNA pseudoknots downstream of the frameshift sites of retroviruses. *Genet. Anal. Technol. Appl. 8*, 191−205.

(8) Du, Z., Giedroc, D. P., and Hoffman, D. W. (1996) Structure of the autoregulatory pseudoknot within the gene 32 messenger RNA of bacteriophages T2 and T6: a model for a possible family of structurally related RNA pseudoknots. *Biochemistry 35*, 4187−4198.

(9) Huang, X., Yang, Y., Wang, G., Cheng, Q., and Du, Z. (2014) Highly conserved RNA pseudoknots at the gag-pol junction of HIV-1 suggest a novel mechanism of −1 ribosomal frameshifting. *RNA 20*, 587−593.
(10) Dinman, J. D., Richter, S., Plant, E. P., Taylor, R. C., Hammell, A. B., and Rana, T. M. (2002) The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc. Natl. Acad. Sci. U.S.A. 99*, 5331−5336.
(11) Dulude, D., Baril, M., and Brakier-Gingras, L. (2002) Characterization of the frameshift stimulatory signal controlling a programmed −1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res. 30*, 5094−5102.
(12) Gaudin, C., Mazauric, M.-H., Traïkia, M., Guittet, E., Yoshizawa, S., and Fourmy, D. (2005) Structure of the RNA signal essential for translational frameshifting in HIV-1. *J. Mol. Biol. 349*, 1024−1035.
(13) Staple, D. W., and Butcher, S. E. (2005) Solution structure and thermodynamic investigation of the HIV-1 frameshift inducing element. *J. Mol. Biol. 349*, 1011−1023.
(14) Baril, M., Dulude, D., Gendron, K., Lemay, G., and Brakier-Gingras, L. (2003) Efficiency of a programmed −1 ribosomal frameshift in the different subtypes of the human immunodeficiency virus type 1 group M. *RNA 9*, 1246−1253.
(15) Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature 460*, 711−716.
(16) Merino, E. J., Wilkinson, K. A., Coughlan, J. L., and Weeks, K. M. (2005) RNA structure analysis at single nucleotide resolution by selective 2′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc. 127*, 4223−4231.
(17) Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.
(18) Deigan, K. E., Li, T. W., Mathews, D. H., and Weeks, K. M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A. 106*, 97−102.
(19) Hajdin, C. E., Bellaousov, S., Huggins, W., Leonard, C. W., Mathews, D. H., and Weeks, K. M. (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Natl. Acad. Sci. U.S.A. 110*, 5498−5503.
(20) Jepsen, J. S., Sørensen, M. D., and Wengel, J. (2004) Locked nucleic acid: a potent nucleic acid analog in therapeutics and biotechnology. *Oligonucleotides 14*, 130−146.
(21) Kaur, H., Babu, B. R., and Maiti, S. (2007) Perspectives on chemistry and therapeutic applications of Locked Nucleic Acid (LNA). *Chem. Rev. 107*, 4672−4697.
(22) Mouzakis, K. D., Lang, A. L., Vander Meulen, K. A., Easterday, P. D., and Butcher, S. E. (2013) HIV-1 frameshift efficiency is primarily determined by the stability of base pairs positioned at the mRNA entrance channel of the ribosome. *Nucleic Acids Res. 41*, 1901−1913.
(23) Chertova, E., Bess, J. W., Crise, B. J., Sowder, R. C., II, Schaden, T. M., Hilburn, J. M., Hoxie, J. A., Benveniste, R. E., Lifson, J. D., Henderson, L. E., and Arthur, L. O. (2002) Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), Is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J. Virol. 76*, 5315−5325.
(24) Adachi, A., Koenig, S., Gendelman, H. E., Daugherty, D., Gattoni-Celli, S., Fauci, A. S., and Martin, M. A. (1987) Productive, persistent infection of human colorectal cell lines with human immunodeficiency virus. *J. Virol. 61*, 209−213.
(25) Means, R. E., Matthews, T., Hoxie, J. A., Malim, M. H., Kodama, T., and Desrosiers, R. C. (2001) Ability of the V3 loop of simian immunodeficiency virus to serve as a target for antibody-mediated neutralization: correlation of neutralization sensitivity, growth in macrophages, and decreased dependence on CD4. *J. Virol. 75*, 3903−3915.
(26) Gherghe, C., Lombo, T., Leonard, C. W., Datta, S. A. K., Bess, J. W., Gorelick, R. J., Rein, A., and Weeks, K. M. (2010) Definition of a high-affinity Gag recognition structure mediating packaging of a retroviral RNA genome. *Proc. Natl. Acad. Sci. U.S.A. 107*, 19248−19253.
(27) Ott, D. E., Coren, L. V., Johnson, D. G., Sowder, R. C., Arthur, L. O., and Henderson, L. E. (1995) Analysis and localization of cyclophilin A found in the virions of human immunodeficiency virus type 1 MN strain. *AIDS Res. Hum. Retroviruses 11*, 1003−1006.
(28) Karabiber, F., McGinnis, J. L., Favorov, O. V., and Weeks, K. M. (2013) QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA 19*, 63−73.
(29) McGinnis, J. L., Duncan, C. D. S., and Weeks, K. M. (2009) High-throughput SHAPE and hydroxyl radical analysis of RNA structure and ribonucleoprotein assembly. *Methods Enzymol. 468*, 67−89.
(30) Mortimer, S. A., and Weeks, K. M. (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc. 129*, 4144−4145.
(31) Steen, K.-A., Rice, G. M., and Weeks, K. M. (2012) Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J. Am. Chem. Soc. 134*, 13160−13163.
(32) Rice, G. M., Leonard, C. W., and Weeks, K. M. (2014) RNA secondary structure modeling at consistent high accuracy using differential SHAPE. *RNA 20*, 846−854.
(33) Blake, R. D., and Delcourt, S. G. (1996) Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res. 24*, 2095−2103.
(34) Marcheschi, R. J., Tonelli, M., Kumar, A., and Butcher, S. E. (2011) Structure of the HIV-1 frameshift site RNA bound to a small molecule inhibitor of viral replication. *ACS Chem. Biol. 6*, 857−864.
(35) Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., Giddings, M. C., and Weeks, K. M. (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol. 6*, e96.
(36) Tyrrell, J., McGinnis, J. L., Weeks, K. M., and Pielak, G. J. (2013) The cellular environment stabilizes adenine riboswitch RNA structure. *Biochemistry 52*, 8777−8785.
(37) Gendron, K., Charbonneau, J., Dulude, D., Heveker, N., Ferbeyre, G., and Brakier-Gingras, L. (2008) The presence of the TAR RNA structure alters the programmed −1 ribosomal frameshift efficiency of the human immunodeficiency virus type 1 (HIV-1) by modifying the rate of translation initiation. *Nucleic Acids Res. 36*, 30−40.
(38) Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., and Lamarre, D. (1996) Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. *J. Virol. 70*, 3763−3769.
(39) Doyon, L., Payant, C., Brakier-Gingras, L., and Lamarre, D. (1998) Novel Gag-Pol frameshift site in human immunodeficiency virus type 1 variants resistant to protease inhibitors. *J. Virol. 72*, 6146−6150.
(40) Girnary, R., King, L., Robinson, L., Elston, R., and Brierley, I. (2007) Structure-function analysis of the ribosomal frameshifting signal of two human immunodeficiency virus type 1 isolates with increased resistance to viral protease inhibitors. *J. Gen. Virol. 88*, 226−235.
(41) Houck-Loomis, B., Durney, M. A., Salguero, C., Shankar, N., Nagle, J. M., Goff, S. P., and D'Souza, V. M. (2011) An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature 480*, 561−564.