Article

Single-Molecule FRET Reveals Three Conformations for the TLS Domain of Brome Mosaic Virus Genome

Mario Vieweger,[1] Erik D. Holmstrom,[1] and David J. Nesbitt[1,*]
[1]Joint Institute for Laboratory Astrophysics, University of Colorado and National Institute of Standards and Technology, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado

ABSTRACT Metabolite-dependent conformational switching in RNA riboswitches is now widely accepted as a critical regulatory mechanism for gene expression in bacterial systems. More recently, similar gene regulation mechanisms have been found to be important for viral systems as well. One of the most abundant and best-studied systems is the tRNA-like structure (TLS) domain, which has been found to occur in many plant viruses spread across numerous genera. In this work, folding dynamics for the TLS domain of Brome Mosaic Virus have been investigated using single-molecule fluorescence resonance energy transfer techniques. In particular, burst fluorescence methods are exploited to observe metal-ion ($[M^{n+}]$)-induced folding in freely diffusing RNA constructs resembling the minimal TLS element of brome mosaic virus RNA3. The results of these experiments reveal a complex equilibrium of at least three distinct populations. A stepwise, or consecutive, thermodynamic model for TLS folding is developed, which is in good agreement with the $[M^{n+}]$-dependent evolution of conformational populations and existing structural information in the literature. Specifically, this folding pathway explains the metal-ion dependent formation of a functional TLS domain from unfolded RNAs via two consecutive steps: 1) hybridization of a long-range stem interaction, followed by 2) formation of a 3′-terminal pseudoknot. These two conformational transitions are well described by stepwise dissociation constants for $[Mg^{2+}]$ ($K_1 = 328 \pm 30$ μM and $K_2 = 1092 \pm 183$ μM) and $[Na^{+}]$ ($K_1 = 74 \pm 6$ mM and $K_2 = 243 \pm 52$ mM)-induced folding. The proposed thermodynamic model is further supported by inhibition studies of the long-range stem interaction using a complementary DNA oligomer, which effectively shifts the dynamic equilibrium toward the unfolded conformation. Implications of this multistep conformational folding mechanism are discussed with regard to regulation of virus replication.

INTRODUCTION

The last three decades of nucleic acid research have revealed a remarkable versatility in RNA functionality (1). This diversity stems from the catalytic activity of noncoding RNA domains, illustrating that the biochemical role of RNA goes far beyond the ability to encode for proteins. Such structural RNA elements are involved in biological catalysis (e.g., self-splicing, ligation) (2,3), posttranscriptional modification (e.g., tRNA maturation, telomere synthesis) (4), and regulation of gene expression (e.g., riboswitches, microRNAs) (5–9). Despite such variability in function, these RNA elements all have one common theme: biological activity depends on efficient, high-fidelity formation of detailed three-dimensional structures (10). Essential in the formation and preservation of active structures are tertiary contacts (11), which are often facilitated by divalent ions and small effector molecules that stabilize the global fold of the RNA (12) against the electrostatic repulsion associated with the negatively charged phosphate backbone (13).

One biological function of noncoding RNAs that heavily relies on these interactions is the regulation of gene expression (14). In conventional gene expression, the functional product, often a protein, is synthesized from messenger RNA (mRNA), which in addition to the coding sequence often contains structural elements within untranslated regions that can be 5′ and/or 3′ to the coding sequence. These untranslated regions offer an excellent way of regulating mRNA recognition by the ribosome and thereby controlling the efficiency of gene expression (15). One exceedingly important group of regulatory RNA elements in gene expression are riboswitches (7), i.e., structural RNA elements found on bacterial mRNAs located upstream of the coding sequence that can generally occupy one of two ligand-dependent conformational states that are either gene expression active (on-state) or inactive (off-state). Modulation of gene expression is achieved through selective stabilization of one or the other conformation by binding to a metabolite. By way of example, a riboswitch that is gene-translation active, i.e., predominantly in the on-state, at low metabolite concentrations, is downregulated at elevated metabolite levels through stabilization of the off-state (16,17). Thus, sensing of the [metabolite] results in self-regulation of gene expression based on the RNAs affinity for metabolite binding.

Bacterial gene regulation by riboswitches represents a highly active area of research, with new classes being

Submitted July 6, 2015, and accepted for publication October 7, 2015.
*Correspondence: djn@colorado.edu
Editor: Lois Pollack.

0006-3495/15/12/2625/12
http://dx.doi.org/10.1016/j.bpj.2015.10.006

discovered at a yearly rate. By way of contrast, similar regulatory mechanisms are only just beginning to emerge for viral systems (18). Viruses are ubiquitous, infectious agents that rely on exploitation of the host cell machinery for replication. The simplest viruses achieve this goal with as few as two functional components: 1) a protective shell made from identical capsid proteins and 2) the viral genome that contains all the information necessary for successful replication. The small virion size (20–100 nm) greatly limits the amount of genetic information that can be encoded within the viral genome (19). As a result, viruses necessitate a particularly efficient and compact system for regulation of their complex replication cycle.

In positive-sense single-stranded RNA ((+)ssRNA) viruses, the genome simultaneously serves as mRNA for expression of viral genes as well as a template for RNA replication. This multifunctional role of (+)ssRNA viral genomes mandates an efficient mechanism for switching from translation to replication as a function of the viral lifecycle. As a result, some viruses have developed specific structural domains at the 3′-end of their genomes to accommodate these additional regulatory requirements (20). One of the most abundant and well-understood classes of viral regulatory domains is the tRNA-like structure (TLS), which is found on many plant viruses spread across many genera (8,9,21). These TLS-domains replace the 3′-poly(A) tail of endogenous mRNA while providing similar functions in translation and stability. Specific virus-related properties of TLS domains include: 1) involvement in encapsidation of viral RNAs, 2) presentation of minus-strand promoter elements for replicase recognition, and 3) regulation of minus-strand synthesis (22). When one considers the intricate structure-function relationship of noncoding RNAs, the three-dimensional structure of this domain clearly plays an essential role in its regulatory mechanism. Interestingly, the TLS domain has been found to bear a remarkable resemblance to the L-shape of cellular tRNAs (Fig. 1 *b*), hence giving it the name TLS or tRNA-like structure. Accordingly, TLS domains retain three key tRNA-like biochemical properties: their abilities to 1) be amino-acylated, 2) form a ternary complex with EF1A translation elongation factors, and 3) serve as substrates for 3′ adenylation by CCA tRNA nucleotidyltransferase. However, the role of such tRNA specific functions in viral replication is not well understood, and indeed varies among viruses (8,9).

A particularly interesting and representative example of TLS domains is found in the Brome Mosaic Virus (BMV) genome. BMV is a plant virus member of the *Bromoviridae* family that is well established as a model system for the infection of (+)ssRNA viruses (24), and has been instrumental for studies of in vitro viral assembly (25). Its genome is divided into three 5′ capped RNAs that encode all four genes required for BMV replication. RNA1 and RNA2 encode for 1) a helicase-like protein and 2) an RNA-dependent RNA polymerase, respectively. These two gene products interact with one another to form the replication complex. RNA3 encodes for two separate proteins, a movement protein that is essential for cell-to-cell migration and a coat protein that is expressed from a subgenomic mRNA (RNA4) produced during the replication cycle. The TLS domains of the three genomic RNAs are nearly identical, with RNA1 and RNA2 differing from RNA3 by only one and three bases, respectively (26). It is therefore somewhat surprising that the three TLS domains show significant differences in their biochemical properties. For example, RNA1 and RNA2 require 3′ amino acylation to support infectivity, but not RNA3. Contrariwise, RNA3 requires the TLS domain as a *cis*-regulating element for effective genome encapsidation, while it can be either *cis*- or *trans*- regulating in RNA1 and RNA2 (8,27).

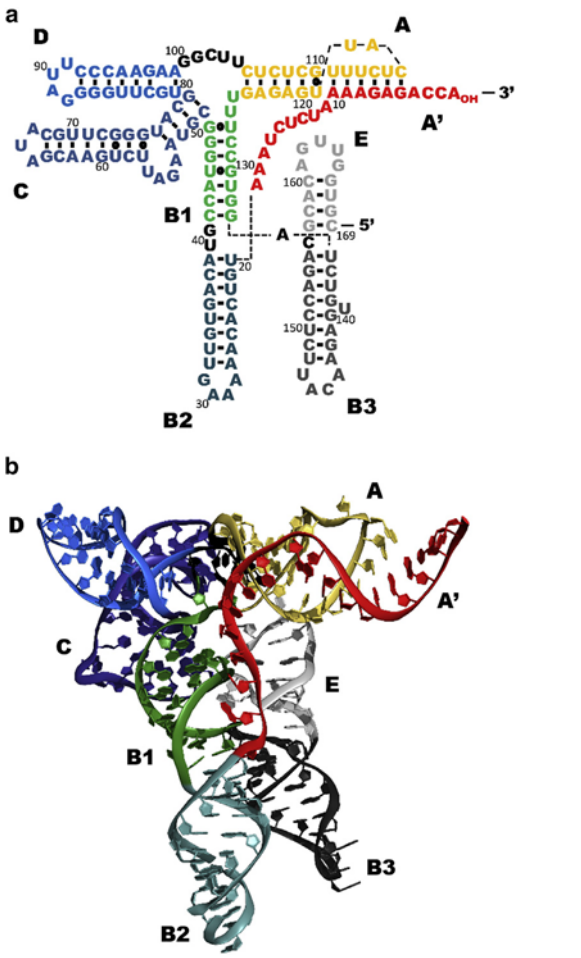

FIGURE 1 TLS domain of Brome Mosaic Virus (BMV). (*a*) Secondary (2°) and (*b*) tertiary (3°) representations of the 3′-terminal 169-nt region of BMV RNA3. To simplify visualization, sections are color-coded consistently between the two representations. Bases are numbered from the 3′-terminus and labeled according to Felden et al. (23). To see this figure in color, go online.

Structural studies involving chemical and enzymatic probing have provided a three-dimensional model of the TLS structure, as illustrated in Fig. 1 *b* (23). While no

atomic resolution structural studies (e.g., NMR and x-ray crystallography) for this TLS motif have been reported, small angle x-ray scattering efforts suggest formation of a compact core similar to the L-shape of tRNA, but with significant flexibility and structural complexity (28,29). These studies reveal two distinct conformations that exist in solutions with either high (i.e., 10 mM) or low (e.g., 0 mM) Mg^{2+},which is consistent with a single two-state [Mg^{2+}]-dependent folding mechanism for the BMV TLS RNA (Fig. 2). At 0 mM [Mg^{2+}], the TLS presents itself as an unfolded (U) string of seven individual stem-loops. Conversely, at 10 mM [Mg^{2+}], the TLS folds into a compact (F) conformation stabilized by two intramolecular interactions (30): 1) a stem interaction between nucleotides 42–48 and 127–133 (Fig. 2, *green*, STEM) and 2) a pseudoknot interaction between nucleotides 5–10 and 111–116 (Fig. 2, *red* and *orange*, PSK). It is worth noting for later discussion that neither extreme of 0 or 10 mM [Mg^{2+}] represents realistic cellular conditions for viral infection.

As a fundamental consequence of this folded structural model, folding from U to F requires formation of two specific interactions in addition to the unzipping of a 5-nt hairpin (Fig. 2 *A′*). This presents a rather complicated scenario for TLS folding, with at least two potential parallel folding pathways differing in the order in which the two intramolecular interactions are formed. One therefore anticipates the population of TLS conformations to be in a complex equilibrium at intermediate, physiologically relevant [Mg^{2+}] values. Detailed dynamics and kinetics studies of such TLS folding pathways as a function of [Mg^{2+}] have not been reported to date, which represent a major goal of this work. Toward this end, we have designed a TLS construct resembling the minimum TLS element of BMV RNA3 to investigate the folding dynamics using single-molecule fluorescence resonance energy transfer methods (31,32). Studies at the single-molecule level permit direct observation of conformational populations (33), providing a particularly powerful tool for investigating structural heterogeneity, intermediate states, and RNA-folding dynamics under equilibrium conditions (34,35). As a result, intermediate or misfolded conformations at the single-molecule level can be explicitly identified and subpopulations readily distinguished (35,36).

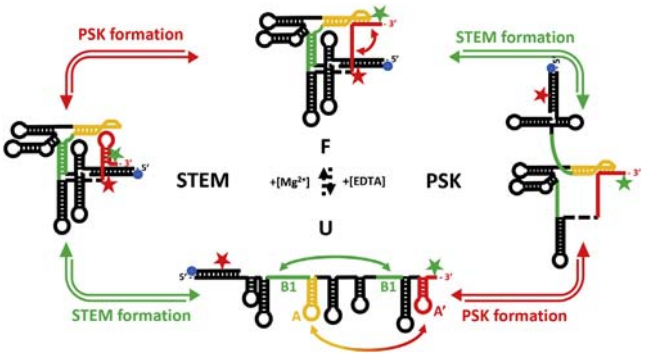

FIGURE 2 Observed TLS dynamics (U ⇄ F) and plausible folding pathways (STEM and PSK intermediates). Two-dimensional secondary structure representations of TLS conformations. Folded (F) and unfolded (U) conformations are based on chemical and enzymatic probing performed at 10 mM [Mg^{2+}] and 1 mM [EDTA], respectively. Folding from U to F conformation involves formation of two intramolecular interactions, 1) a stem (*green arrow* labeled STEM formation) and 2) a pseudoknot (*red arrow* labeled PSK formation). STEM and PSK represent intermediate conformations based on stepwise folding pathways differing in the order of formation of the intramolecular interactions. See text for details. To see this figure in color, go online.

In this work, single-molecule burst fluorescence methods are utilized to experimentally observe the metal-ion ([M^{n+}])-dependent folding of the BMV TLS domain under freely diffusing conditions by exploiting FRET between a Cy3 donor and Cy5 acceptor covalently attached near the 3′ and 5′ termini, respectively (Fig. 2). TLS folding resulting from formation of intramolecular interactions is observed as a change in the distance (*R*) between the termini as measured by changes in FRET efficiencies (E_{FRET}). Histograms of E_{FRET} probabilities reveal three distinct conformations of TLS molecules in solution. Evolution of the fractional populations with [M^{n+}] is found to be in excellent agreement with a sequential stepwise folding model based on contemporary structural information (Figs. 1 and 2). Most importantly, this model requires folding through a stable but never-before reported intermediate state, which we temporarily denote by I. As we hope to make a convincing case for in this work, this, to our knowledge, new conformational state contains the compact core stabilized by the STEM interaction and the 3′ hairpin A′, as determined by oligo competition studies designed to inhibit formation of the STEM domain. Finally, the implications of such a folding mechanism as they relate to regulation of the viral replication cycle are discussed.

MATERIALS AND METHODS

TLS construct synthesis

To best retain biological function, the RNA construct is designed to closely resemble the 3′ terminal TLS of BMV RNA3, which contains the minimum length element of 169 nt (37). In an effort to probe [Mg^{2+}]-dependent RNA folding on a single-molecule level, three modifications are incorporated: 1) a Cy3/Cy5-FRET pair (*green* and *red stars*, Fig. 2) for observation of conformational dynamics and 2) a biotin-anchor (*blue circle*, Fig. 2) to enable studies of immobilized molecules. Briefly, construct synthesis is achieved by: 1) synthesis of a DNA template containing the minimum TLS element (169 nt), a 5′ extension (20 nt), and T7 promoter sequence (25 nt) using Assembly PCR; 2) in vitro transcription of the DNA template (214 nt) using T7 RNA polymerase followed by gel extraction of the structural RNA element (189 nt); 3) selective 3′ end-labeling (38) via oxidation of the 3′-ribose to dialdehyde with sodium meta-periodate ($NaIO_4$), followed by hydrazone formation by reaction of the aldehyde with a hydrazide-activated Cy3 fluorophore; 4) reductive amination of the hydrazone bond with sodium cyanoborohydride ($NaCNBH_3$); and 5) annealing of the 3′-labeled RNA construct to a complementary oligomer containing a Cy5 label and Biotin modification for immobilization (tether,

5′-UGUGU/iCy5/GUGUGUGUGUGCG GGCCC /BioTEG/-3′; Integrated DNA Technologies, Coralville, IA) followed by reverse-phase HPLC purification of doubly labeled RNA.

Sample preparation

Samples for single-molecule studies are prepared as dilute (~125 pM) RNA solutions in 50 mM HEPES buffer (H9897; Sigma Aldrich, St. Louis, MO) at pH 7.5 containing a fluorescence support system comprised of the triplet quencher Trolox (2 mM, 9-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid, 238813; Sigma Aldrich) and the enzymatic oxygen scavenging system PCA/PCD (10 mM, 3,4-dihydroxybenzoic acid, 37580; 0.1 mg/mL, protocatechuate 3,4-dioxygenase, P2879; Sigma Aldrich) to increase the photostability of the Cy3/Cy5 FRET pair (39,40). Imaging solutions are studied in a microliter flowcell assembled from a microscope slide and a No. 1.5 cover glass (22 × 22 mm; Corning, Corning, NY) secured by two pieces of double-sided tape separated by 1–2 mm to form a narrow channel with a ~5 μL sample volume. This channel is flushed with 50 μL of sample solution just before the experiment. Salt-dependent folding-unfolding equilibrium studies in BMV TLS are performed by systematic addition of concentrations of NaCl or $MgCl_2$ solution. Similarly, competition studies are performed by addition of DNA oligomers complementary to the STEM domain at a series of known concentrations.

Single-molecule instrumentation

Single-molecule experiments are performed on an inverted scanning confocal microscope (IX70; Olympus, Center Valley, PA) with dual laser excitation, displayed in Fig. 3 *a* (labels in *bold* refer to Fig. 3 *a*) (41,42). Briefly, pulses from a mode-locked, frequency doubled Nd:YAG laser (EX1, 10 ps, 532 nm, LYNX; Time-Bandwidth Products, Zurich, Switzerland) are used to trigger synchronized pulses from a diode laser (EX2, < 100 ps, 635 nm, PDL800; PicoQuant, Berlin, Germany) at a repetition rate of 20.6 MHz. This pair of lasers is then spatially overlapped (DM1) and temporally delayed to obtain an interleaved excitation train of uniformly spaced, alternating red/green laser pulses separated by 25 ns. The laser beams are expanded (BE) to overfill the back aperture of the objective and focused to the diffraction limit by a 1.2 NA water immersion objective (×60W, UPlanApo 60×; Olympus). Intensities are attenuated by neutral density filters to deliver 0.5–100 μW at the focal plane, with emitted photons collected by the same objective and filtered from incident laser light with a dichroic mirror (DM2, ZT532/640rpc; Chroma Technology, Bellows Falls, VT). Out-of-plane and off-axis fluorescent photons are spatially filtered by a 50-μm confocal pinhole (P). The beam of fluorescent photons is split into horizontally and vertically polarized detection paths (PBS, CVI PBSH-450-1300-100; CVI Laser Optics, Albuquerque, NM), which are further separated into donor (CY3) and acceptor (CY5) channels by identical dichroic mirrors (DM3, 645DCXR; Chroma Technology). The four photon channels are band-pass-filtered (HQ585/70M and HQ700/75M; Chroma Technology) and detected on individual single-photon counting avalanche photodiodes (APDs, SPCM-AQR-14; PerkinElmer Optoelectronics, Fremont, CA). Photon arrival events are recorded on a time-correlated single-photon counting module (SPC-134; Becker & Hickl, Berlin, Germany) and stored time-tagged and time-resolved with respect to the start of the experiment and the excitation pulse train.

Analysis of FRET efficiencies from freely diffusing molecules

Analysis of freely diffusing molecules has been described in detail in Fiore et al. (41). Briefly, fluorescence from labeled RNA constructs is collected with the laser focus 15 μm into the sample solution to ensure that experimental observations are not influenced by the presence of the flowcell surface. As a molecule traverses through the confocal volume (~0.5 fL), it can absorb and emit photons several times, resulting in a fluorescence burst, which is then separated spatially with respect to emission from donor and acceptor molecules and detected as time-tagged and time-resolved with respect to the alternating excitation train (alternating laser excitation, ALEX). Time traces of fluorescence bursts are obtained for both donor (Fig. 3 *b*) and acceptor excitation (Fig. 3 *c*) by sorting single-photon arrivals into 1-ms bins. Individual bursts are identified as single time-bins that exceed a count-rate threshold (typically 25 kHz) as a result of fluorescence from either of the two alternating excitation sources (33) to ensure that only doubly labeled RNA constructs are considered for analysis. For successful events, FRET efficiencies (E_{FRET}) are calculated from corrected donor (I^c_D) and acceptor (I^c_A) intensities according to Eq. 1 to compensate for molecular (direct excitation, differential quantum yields) and instrumentation (collection efficiencies, channel cross talk) parameters (41):

$$E_{FRET} = \frac{I^c_A}{I^c_A + I^c_D}. \quad (1)$$

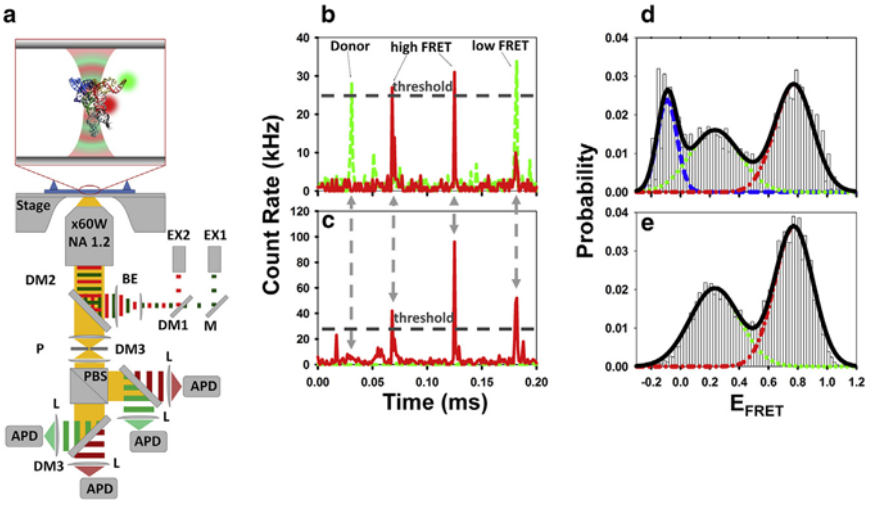

FIGURE 3 Experimental techniques. (*a*) Confocal microscope for single-molecule studies (see text for details). Freely diffusing molecules are alternately excited at the donor and acceptor wavelengths. The fluorescence photon is detected on one of four avalanche photodiodes, sorted by (donor/acceptor) wavelength and (horizontal/vertical) polarization. (*b*) Time-trace of fluorescence intensities for freely diffusing TLS molecules with donor excitation. The dashed trace represents emission from donor molecules, while the solid trace represents emission from acceptor molecules excited by energy transfer. (*c*) Time-trace of fluorescence intensities for freely diffusing TLS molecules excited directly on Cy5, which unambiguously confirms the presence of the acceptor dye and permits ALEX filtering. (*d*) Histogram of fluorescence bursts from intensity traces of donor-excited molecules that reach the intensity threshold, where the distribution at negative FRET efficiencies originates from donor-only molecules due to rigorous correction for both cross talk and direct excitation of Cy5. (*e*) Histogram of FRET efficiencies from fluorescence bursts that exceed the threshold following donor and acceptor excitation (ALEX-filtered). To see this figure in color, go online.

Energy transfer efficiencies are then histogrammed for ~10^3–10^4 individual events, with distributions analyzed by fitting to a sum of Gaussians to determine the statistical probabilities of observing RNA molecules with a given E_{FRET}. As shown in Fig. 3, *d* and *e*, this ALEX filtering procedure permits efficient removal of the donor-only contribution at negative E_{FRET} values (Fig. 3 *d*) without distorting the E_{FRET} distributions.

RESULTS AND DISCUSSION

TLS folding (U ⇌ F) occurs through an intermediate conformation

To correlate structural information from chemical and enzymatic probing studies (U and F in Fig. 2) with E_{FRET} signatures in single-molecule experiments, burst fluorescence methods are utilized to observe [M^{n+}]-induced folding of the BMV TLS domain in freely diffusing molecules (31,32). Placement of a Cy3 donor and Cy5 acceptor near the 3′ and 5′ termini (Fig. 2) allows observation of U to F folding as measurable change in FRET efficiencies determined according to Eq. 1. In agreement with structure probing studies, [Mg^{2+}]-dependent TLS folding can be systematically induced by increasing the [Mg^{2+}] from 0 to 10 mM under standard HEPES buffer conditions (50 mM HEPES, 50 mM [Na^+], pH 7.5). Representative histograms are displayed in Fig. 4 *a* for three experimental conditions, demonstrating the [Mg^{2+}]-dependent distribution of E_{FRET}.

In clear contrast with the accepted folding model (Fig. 2), three rather than two populations are distinctly observed, evolving with increasing [Mg^{2+}] from 1) low to 2) high, and then back down to 3) an intermediate E_{FRET} value. As the titration is performed over a narrow range in total ionic strength, structural rearrangements within subpopulations are minimal, permitting analysis as a three-state system assuming well-defined and distinguishable distributions (41). Consequently, the entire set of E_{FRET} histograms can be fit simultaneously to a sum of three Gaussians with common but adjustable peak widths and centers, to obtain an accurate determination of subpopulations across the entire range of investigated [Mg^{2+}]. The resulting least-squares fits clearly support the presence of three conformational subpopulations (Fig. 4 *a*) at E_{FRET} = 0.097 ± 0.002 (*dashed*), 0.744 ± 0.006 (*dashed-dotted*), and 0.311 ± 0.006 (*dotted*), in order of increasing [Mg^{2+}].

The presence of a stable, unreported RNA conformation at high E_{FRET} and indeed more physiologically relevant [Mg^{2+}] represents an interesting observation that warrants further discussion. One key question is why this conformation was unobserved in previous structure-probing studies The simple answer is that these previous ensemble studies were intentionally performed at 1) low [Mg^{2+}] (50 mM HEPES, 1 mM EDTA) and 2) high [Mg^{2+}] (50 mM HEPES, 10 mM Mg^{2+}) conditions (23), precisely to achieve more nearly monodisperse conformational populations. Indeed, these low and high [Mg^{2+}] conditions do yield primarily unfolded (U) and folded (F) species (see Fig. 2). On the other hand, single-molecule burst fluorescence methods achieve higher sensitivity to conformational change (31–34,42,43) and subpopulations (33), thus providing valuable kinetic and thermodynamic information on the folding pathway from U to F. Correlation of structural and single-molecule information suggests that populations at low and intermediate E_{FRET} (*dashed* and *dotted curves* in Fig. 4 *a*) correspond to the U and F conformations in Fig. 2, respectively. The population at high E_{FRET} (*dashed-dotted curve*, Fig. 4 *a*) features a shorter distance between the FRET pair than either U and F, and thus represents a predominant albeit heretofore unobserved intermediate species (I) along the U to F folding pathway under physiological [Mg^{2+}] conditions. Further validation for this assignment of TLS conformations to E_{FRET} populations can be inferred from the known distance dependence for FRET efficiencies. Based on helical constraints, the U conformation, consisting of a string of stem loops, is expected to exhibit a low E_{FRET} signature with a distribution centered ~0 (> 10 nm 5′ to 3′ distance) at full extension. Due to residual monovalent salt present under U conditions, an electrostatically equilibrated structure is anticipated, consistent with the observed

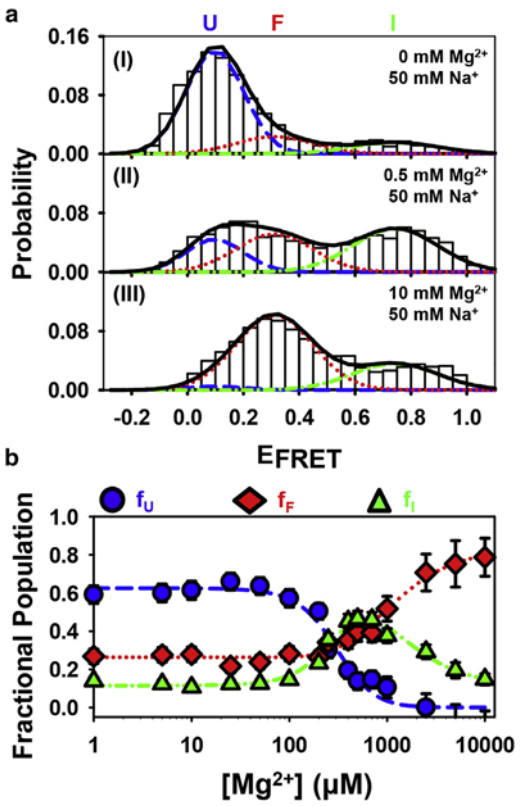

FIGURE 4 [Mg^{2+}]-induced TLS folding. (*a*) Histograms of E_{FRET} distributions at (I) low, (II) intermediate, and (III) high [Mg^{2+}] at 50 mM [Na^+]. The black solid curve is fit to a sum of three Gaussians with the individual distributions displayed in dashed, dotted, and dashed-dotted, respectively. (*b*) Fractional populations from Gaussian distributions are plotted versus [Mg^{2+}]. Solid curves are fits to a three-state stepwise folding mechanism (see text for details). To see this figure in color, go online.

E_{FRET} of ~0.1. By way of contrast, the F conformation forms a compact core stabilized by STEM and PSK intramolecular interactions (see Fig. 1, *top* and *bottom*), effectively reducing the 5′–3′ distance, thus increasing the FRET efficiency to E_{FRET} ~0.3–0.6 (see Fig. 2), i.e., again consistent with the observed E_{FRET} value of 0.311(6).

The high $E_{FRET} = 0.744(6)$ value for the intermediate (I) originates from a conformation between U and F with a short donor-acceptor distance, indicative of a stepwise folding pathway that favors one of the two possible pseudoknot (PSK) or stem (STEM) interactions. Unfolding of the folded F conformation through a PSK intermediate (Fig. 2) would require rotation of the 5′ terminus about the structural core (Fig. 1 *b*), likely resulting in further separation of the donor-acceptor distance (23), thus predicting an E_{FRET} value between those observed for F and U. By way of contrast, unfolding of F through a STEM intermediate would likely proceed via unzipping the 3′ PSK interaction and formation of hairpin A′ (Fig. 2), likely resulting in a reduced 3′-5′ distance and thus increasing E_{FRET}, as observed in this single-molecule experiments. Indeed, from structure modeling studies, Felden et al. (23) have even postulated the existence of a mostly folded conformation where hairpin A′ is formed as opposed to the pseudoknot in the F conformer (23). Consequently, we propose a TLS folding pathway from a fully unfolded (U) to fully folded (F) structure that proceeds via formation of the STEM interaction followed by the 3′ pseudoknot (PSK). This implies the presence of a stable STEM conformation for TLS that predominates at physiologically intermediate [M^{n+}] concentrations (see STEM pathway in Fig. 2).

TLS folding follows a stepwise mechanism (U # I # F)

To further validate the presence of a three-state stepwise folding mechanism, we analyze the data in the context of the following simple equilibrium kinetic model:

$$U \# I \quad K_1;$$

$$I \# F \quad K_2:$$

For such a model, K_1 and K_2 represent Mg^{2+}-dependent dissociation constants for the first (U # I) and second (I # F) steps, respectively. Two approaches are commonly used to describe such M^{n+}-dependent RNA folding equilibria:mass action schemes and formalisms based on interaction coefficients. While mass action schemes assume that ions are ligands that bind to RNA sites in a stoichiometric manner, interaction coefficients consider the electrostatic character of interactions between charged macromolecules and ions and thus account for long-range electrostatic interactions and effects of counterions (44). Both approaches ultimately model a sigmoidal response to M^{n+} folding and thus capture the curve of the folding transition and in particular, the midpoint. To demonstrate the stepwise equilibrium process in TLS folding, the more simplistic approach of deriving kinetic models based on stoichiometric mass action schemes is applied. This gives the flexibility to test the validity of the stepwise folding model by further including competitive inhibition of the intermediate state I (see Stepwise TLS Folding Occurs through a STEM Rather than PSK Intermediate). From [M^{n+}]-induced RNA folding studies, it can be readily shown that fractional populations for such a stepwise model should be given by

$$f_i = \frac{[i]}{[U] + [I] + [F]} = \frac{[i]}{1 + \frac{[M]^n}{K_1^n} + \frac{[M]^{n+m}}{K_1^n K_2^m}}; \qquad (2)$$

where n and m values are the degree of Hill cooperativity for Mg^{2+} binding. Experimentally, $f(N_i)$ is proportional to the fractional number of FRET events in each population subgroup,

$$f(N_i) = [N_i]/([N_U] + [N_I] + [N_F]);$$

where N_i is obtained by integrating over each Gaussian lineshape in the single-molecule burst distributions with i = U, I, or F representing the conformation of interest. The [Mg^{2+}]-dependent evolution of these populations is plotted in Fig. 4 *b* (*blue circles* for f_U, *red diamonds* for f_F, and *green triangles* for f_I). As expected, U dominates (63%) at low [Mg^{2+}], although there are also appreciable amounts of TLS in intermediate (12%) and folded (26%) conformations. With increasing [Mg^{2+}], the fractional populations yield a transient peak in the I conformation at near-physiological concentrations (~500 μM). At even higher [Mg^{2+}] (1–10 mM), [I] drops in favor of F state formation down to a small but finite value.

For the simplified kinetic model in Eq. 2, the low and high Mg^{2+} limits should yield entirely unfolded (U) or folded (F) populations, i.e., close to but not observed in Fig. 4. All three states are observed in dynamic equilibrium over the entire range of [M^{n+}] with the unfolded U and folded F conformations dominating at low and high [Mg^{2+}], respectively. Nonetheless, [Mg^{2+}] is not the only parameter affecting the RNA folding equilibrium; other influences, e.g., monovalent ions [M^+] and temperature, visibly affect the folding dynamics resulting in deviations from Eq. 2 in the form of saturating upper and lower boundaries of conformational populations and thus equilibrium constants. Incomplete unfolding ($f_U < 1$) at 0 mM [Mg^{2+}] can thus be ascribed to residual monovalent [Na^+] in the imaging buffer (50 mM HEPES, 50 mM [Na^+]), while incomplete folding ($f_F < 1$) at 10 mM [Mg^{2+}] is most likely due to saturation of the equilibrium (i.e., $K_{eq} \sim 4$ for I # F in Fig. 4 *b*).

Consistent with previous studies, these experimental parameters can be taken into account by offsets f_{pre} and f_{sat} to capture the fraction of prefolded molecules and saturation of the equilibrium due to applied experimental conditions, respectively, resulting in a reduced fraction for the fractional populations (f_i):

$$f_i = f_{pre;i} + f_{pre;i} - f_{sat;i} \frac{[i]}{1 + \frac{[M]^n}{K_1^n} + \frac{[M]^{n+m}}{K_1^n K_2^m}}: \quad (3)$$

The fractional populations $f(N_i)$ have been fit simultaneously to Eq. 3, keeping K_1, K_2, n, and m adjustable but common for all three data sets (f_U, f_I, and f_F). Results are displayed in Table 1, with the least-squares fits overlaid in Fig. 4 *b*. The stepwise folding dynamics are well described by $K_1 = 328 \pm 30$ μM and $n = 2.0 \pm 0.3$ and $K_2 = 1092 \pm 183$ μM and $m = 1.3 \pm 0.2$ for the first and second steps, respectively, yielding a total cooperativity of $n + m = 3.3 \pm 0.4$.

The min/max (f_{pre}, f_{sat}) and dynamic range (Δf) for each conformation provides additional information about the TLS folding dynamics. For example, $f_{sat,U} = 0.63 \pm 0.02$ at $[Mg^{2+}] = 0$ mM implies that even at 50 mM $[Na^+]$ in the buffer, 37% of the TLS molecules are already either fully folded (F) or in the intermediate (I) state. Furthermore, of these 37%, the ratio is $f_{pre,F} = 0.26 \pm 0.01 / f_{pre,I} = 0.12 \pm 0.02 = 2.2{:}1$, which implies a lower free energy for the folded versus intermediate state at 0 mM $[Mg^{2+}]$.

Monovalent metal-ions [M^{Δ}] affect the TLS folding equilibrium

The above $[Mg^{2+}]$-dependent studies clearly indicate a reduced albeit finite sensitivity to monovalent ions $[M^+]$. To assess the importance of such $[M^+]$-induced effects, TLS folding has also been studied as a function of $[Na^+]$, holding $[Mg^{2+}] = 0$ mM by addition of 1 mM [EDTA]. Due to residual concentration of monovalent ions in the imaging buffer solutions (e.g., 50 mM hemi-sodium HEPES at pH 7.5 contains 25 mM $[Na^+]$), the lowest $[Na^+]$ concentration attainable in our $[M^+]$ titration experiments is 35 mM. Burst experiments are therefore performed by increasing $[Na^+]$ from 35 mM to 1.035 M, with representative histograms for 1) low (35 mM), 2) intermediate (135 mM), and 3) high (1035 mM) $[Na^+]$ presented in Fig. 5 *a*. Once again, three distinct populations (U, I, and F) are observed, with populations that evolve in predominance from low through high to intermediate E_{FRET} with increasing $[Na^+]$. Similar to the analysis for the $[Mg^{2+}]$ series, the E_{FRET} distributions have been globally fit to a sum of three Gaussians with adjustable peak widths and centers, resulting in distributions centered at $E_{FRET} = 0.078 \pm 0.002$ (*dashed*, U), 0.718 ± 0.004 (*dashed-dotted*, I), and 0.297 ± 0.005 (*dotted*, F). These E_{FRET} center values are quite close to but lower to those observed in the $[Mg^{2+}]$ titrations ($\Delta E_{FRET}(U) = -0.019 \pm 0.003$, $\Delta E_{FRET}(I) = -0.026 \pm 0.007$ and $\Delta E_{FRET}(F) = -0.015 \pm 0.008$), indicating 1) a slightly more compact overall conformation with $[Mg^{2+}]$ present and 2) shorter distances between the FRET donor and acceptor.

Fractional populations for each of the three component conformations are plotted versus $[Na^+]$, with least-squares fits to Eq. 3 displayed in Table 2 and overlaid in Fig. 5 *b*. The evolution of the equilibrium E_{FRET} populations is again quite well described by a sequential three-state stepwise

TABLE 1 Least-squares fit parameters for [$Mg^{2\Delta}$]-induced TLS conformational change (U / I, I / F) based on an equilibrium 3-state sequential folding model

Parameter	*U*	*I*	*F*
f_{pre}	0.00 ± 0.02	0.12 ± 0.02	0.26 ± 0.01
f_{sat}	0.63 ± 0.02	0.73 ± 0.07	0.83 ± 0.03
Δf ($=f_{sat}$-f_{pre})	0.63 ± 0.03	0.61 ± 0.07	0.57 ± 0.04
K_{D1} (μM)		328 ± 30	
K_{D2} (μM)		1092 ± 183	
n		2.0 ± 0.3	
m		1.3 ± 0.2	

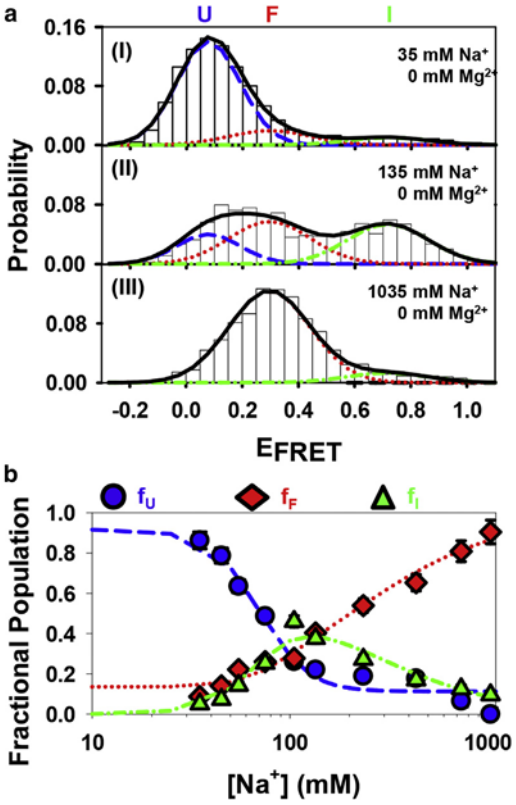

FIGURE 5 $[Na^+]$-induced TLS folding. (*a*) Histograms of E_{FRET} distributions at (I) low, (II) intermediate, and (III) high $[Na^+]$ at 0 mM $[Mg^{2+}]$. Black solid curves are first to a sum of three Gaussians with individual distributions displayed in dashed, dotted, and dashed-dotted, respectively. (*b*) Fractional populations from Gaussian distributions are plotted versus $[Na^+]$. Curves are fits to a three-state stepwise folding mechanism (see text for details). To see this figure in color, go online.

TABLE 2 Least-squares fit parameters for [$Na^{\square}$]-induced TLS conformational change (U / I, I / F) based on an equilibrium 3-state sequential folding model

Parameter	*U*	*I*	*F*
f_{pre}	0.11 ± 0.03	0.00 ± 0.06	0.14 ± 0.04
f_{sat}	0.92 ± 0.08	0.62 ± 0.10	0.98 ± 0.09
$\square f (= f_{sat}\text{-}f_{pre})$	0.81 ± 0.09	0.62 ± 0.11	0.84 ± 0.10
K_{D1} (μM)		74 ± 6	
K_{D2} (μM)		243 ± 52	
n		3.3 ± 0.8	
m		1.3 ± 0.3	

mechanism, with dissociation and coordination constants of $K_1 = 74 \pm 6$ mM, $n = 3.3 \pm 0.8$ and $K_2 = 243 \pm 52$ mM, $m = 1.3 \pm 0.3$ for the U # I and I # F folding steps, respectively. Note that folding is significantly less sensitive to the presence of monovalent versus divalent cations, indeed requiring 225 ± 28- and 222 ± 60-fold higher $[Na^+]/[Mg^{2+}]$ to achieve formation of the PSK and STEM interactions. However, such behavior is in good qualitative agreement with known efficiencies for RNA folding induced by $[Mg^{2+}]$ versus $[Na^+]$, and which has been interpreted in terms of reduced screening of electrostatic repulsion by monovalent versus divalent species (45). More quantitatively, $[Mg^{2+}]$ in $[M^{n+}]$-induced folding of the tetraloop receptor interaction has been observed to be ~500-fold more efficient than $[Na^+]$ (41).

From Fig. 5, the $[Na^+]$ titration data predicts a saturation limit of $f_{sat,U} = 0.92 \pm 0.08$ for the fraction of unfolded molecules in the absence of metal ions, i.e., indistinguishable from unity. A contribution due to the folded fraction of $f_{pre,F} = 0.14 \pm 0.04$ is observed, while the intermediate population ($f_{pre,I} = 0.00 \pm 0.06$) is essentially nonexistent. The saturation limit for formation of the fully folded state ($f_{sat,F} = 0.98 \pm 0.09$) is essentially unity. These results indicate that 1) the same intermediate is formed in both Na^+ and Mg^{2+} folding studies and 2) all molecules are capable of forming both intramolecular interactions at high enough monovalent ion concentrations.

Stepwise TLS folding occurs through a STEM rather than PSK intermediate

To verify that the high E_{FRET} populations in Figs. 4 and 5 reflect a STEM intermediate and thus confirm the predominance of the STEM folding pathway, we take advantage of the fact that the intermediate I likely contains only the PSK or STEM interactions and thus can be selectively influenced by oligo competition studies. Specifically, an 8-nt DNA oligomer (oligo; 5′-CGGTACCC-3′, Integrated DNA Technologies) complementary to nucleotides 41–48 (5′-GGGUACCG-3′) in the STEM domain (*green* in Fig. 2) is utilized in burst fluorescence studies. Incorporation of this oligomer into the TLS construct results in formation of an 8-basepair helix in place of the STEM domain, which is expected to still allow the formation of the 3′ terminal PSK interaction. As seen in Fig. 1, the 5′ domains including B1 hinge on hairpin A′ via a flexible UU linker that facilitates unzipping of the STEM domain without interfering with the 3′ terminal PSK interaction. Oligomer binding will therefore inhibit formation of the STEM conformation, thus shifting the equilibrium in favor of the U conformation. As a consequence, the fractional population of the intermediate conformation (f_I) should decrease with increasing [oligo] for the STEM pathway, while no change is anticipated for the PSK pathway, thus allowing clear differentiation between the PSK and STEM intermediates based on response to oligomer competition. Furthermore, nearest-neighbor calculations predict a dissociation constant (K_c) for competitor binding of $K_c \sim 10$ nM at 20°C, which is readily attainable in burst titration experiments (46).

To modify the three-state stepwise folding equilibrium mechanism to account for inhibition via the complementary oligo, expressions for the fractional populations of the intermediate state at high E_{FRET} can be derived for both PSK and STEM pathways, resulting in Eqs. 4 and 5, respectively:

$$f_{c;STEM} = \frac{\frac{[M]^n}{K_1^n}}{\frac{[oligo]}{K_c} + 1 + \frac{[M]^n}{K_1^n} + \frac{[M]^{n+m}}{K_1^n K_2^m}}; \quad (4)$$

$$f_{c;PSK} = \frac{\frac{[M]^n}{K_1^n} + \frac{[M]^n[oligo]}{K_1^n K_c}}{\frac{[oligo]}{K_c} + 1 + \frac{[M]^n}{K_1^n} + \frac{[M]^n[oligo]}{K_1^n K_c} + \frac{[M]^{n+m}}{K_1^n K_2^m}}: \quad (5)$$

As expected, Eqs. 4 and 5 for the STEM and PSK pathways predict a decrease and increase, respectively, for $f_{c,STEM}$ with increasing [oligo], which should be readily distinguishable.

To test these predictions, burst fluorescence experiments have been performed at buffer conditions favoring the intermediate conformation (I) at high E_{FRET} (135 mM $[Na^+]$, 50 mM HEPES buffer) for [oligo] ranging from 0 to 2.5 μM. E_{FRET} histograms show a systematic decrease in the I state population at high E_{FRET}, as indeed expected for the STEM pathway. Fractional populations are determined from these burst studies by fitting E_{FRET} distributions with a sum of two Gaussians according to $f(I) = [\text{high } E_{FRET}]/([\text{low } E_{FRET}]+[\text{high } E_{FRET}])$ and plotted against [oligo] in Fig. 6 *b*. To more quantitatively establish the STEM conformation as the intermediate for the folding pathway, fractional high and low E_{FRET} populations are fit simultaneously to Eq. 4, fixing all dissociation/cooperativity parameters at values obtained under [oligo] = 0 nM conditions. Least-squares fit predictions to the burst fluorescence data are included (*dashed curve*) in Fig. 6 *b*. Note the very

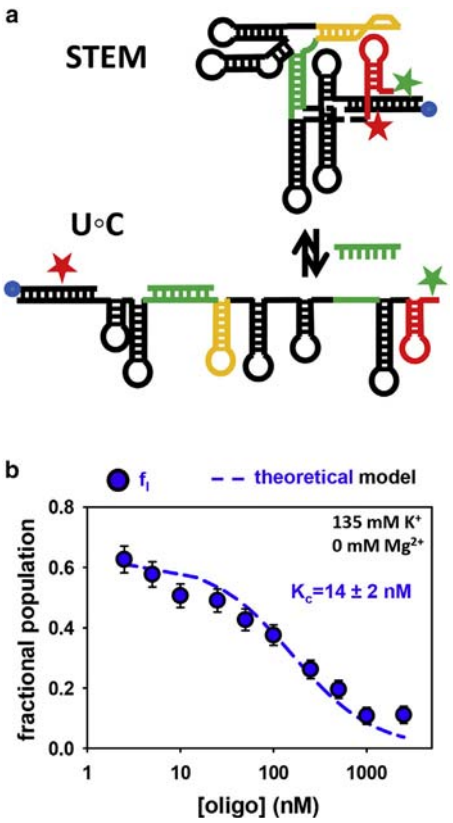

FIGURE 6 Folding inhibition experiments. (*a*) Folding model for competitive inhibition of STEM interaction. Experiments are performed at conditions that favor the intermediate conformation (135 mM [K^+], 0 mM [Mg^{2+}]). Complementary oligomer binds to the unfolded conformation resulting in a complexed state U◦C. (*b*) Evolution of the intermediate conformation f_I is plotted versus [oligo]. The fractional population is fit to a model for competitive inhibition in the STEM pathway (*blue dashed line*). The least-square fitted dissociation constant for the oligomer-RNA binding step ($K_c = 14 \pm 2$ nM) compares well to the predicted $K_D = 10$ nM value. To see this figure in color, go online.

clear agreement between experiment and the oligomer competition model for three-state stepwise TLS folding with a STEM intermediate, with the measured inhibition constant ($K_c = 14 \pm 2$ nM) even in close agreement with estimates ($K_D = 10$ nM) from nearest-neighbor predictions. Indeed, these oligo competition experiments provide unambiguous identification of the I intermediate in a sequential three-state TLS folding mechanism to be the STEM rather than the PSK conformation.

Biological implications

Physiological conditions favor the I (STEM) # F equilibrium step

The above [Mg^{2+}] and [Na^+] titration experiments make clear that a distribution of U, STEM, and F populations is present in dynamic equilibrium. Thus the observed values of $K_1 = 74 \pm 6$ mM and $K_2 = 243 \pm 52$ mM in the [Na^+] series and $K_1 = 328 \pm 30$ μM and $K_2 = 1092 \pm 183$ μM in the [Mg^{2+}] series constitute approximate values for experiments performed in the presence of near-physiological concentrations of both mono- and divalent ions. Previously, presence of both mono- and divalent ions have been reported to exert synergistic as well as competitive effects on the folding dynamics for different RNA constructs, thus predicting opposite effects on the overall dissociation constants (K_D) for the individual steps in such a sequential mechanism (41,47). By way of specific example, the analogous study of [Mg^{2+}]-dependent TLS folding at 135 mM [Na^+] reveals that $K_2 \sim 550$ μM (data not shown), i.e., twofold lower than observed at 35 mM [Na^+]. The effects of mono- and divalent ion contributions on TLS folding are thus synergistic for the K_2 equilibrium step and further shift the midpoint of the folding transitions toward lower [M^{n+}] additionally favoring the fully folded conformer. Most importantly, under physiological conditions, the equilibrium populations will thus be dominated by the second process (STEM # F), while the unfolded (U) state does not contribute significantly.

TLS domains in the BMV genome differ only in the STEM interaction region

As a consequence of this equilibrium, the thermodynamic stability of the STEM intermediate is anticipated to be particularly relevant to BMV biochemical function. While an almost complete sequence identity is observed across the TLS domains of the three genomic RNAs, intriguingly, the few differences that do exist are located within the 7-bp overlap of the long-range stem interaction (26) that was identified above as the sole structure motif responsible for stabilizing the intermediate STEM conformation. As displayed in Fig. 7 *a*, this stem region in RNA3 folds via relatively strong interaction of five canonical (4× GC and 1× AU) and 2 noncanonical GU basepairs, and with nearest-neighbor calculations predicting a free energy change of $\Delta G^0 = -10.3$ kcal/mol. In BMV2, a mutation at A_{44} to G destabilizes the STEM interaction by replacement of an AU with a GU basepair, with the corresponding interaction free energy dropping to $\Delta G^0 = -9.3$ kcal/mol. BMV1 exhibits two additional mutations (U_{131} to C and C_{43} to U) that compensate each other through replacement of a GU by a GC and a GC by a GU basepair resulting in $\Delta G^0 = -9.2$ kcal/mol. Consequently, the STEM interaction is weakened in BMV1 and BMV2 by approximately the same amount ($\Delta\Delta G^0 = 1.05 \pm 0.05$ kcal/mol) with respect to BMV3 effectively reducing the thermodynamic stability of their STEM intermediates. Considering the stepwise TLS folding mechanism in Fig. 2, formation of the long-range stem and pseudoknot interactions are in a thermodynamic competition for the predominate folding intermediate. While the STEM intermediate outperforms the PSK in BMV3, weakening of the stem interaction in BMV1 and BMV2 would likely shift the equilibrium

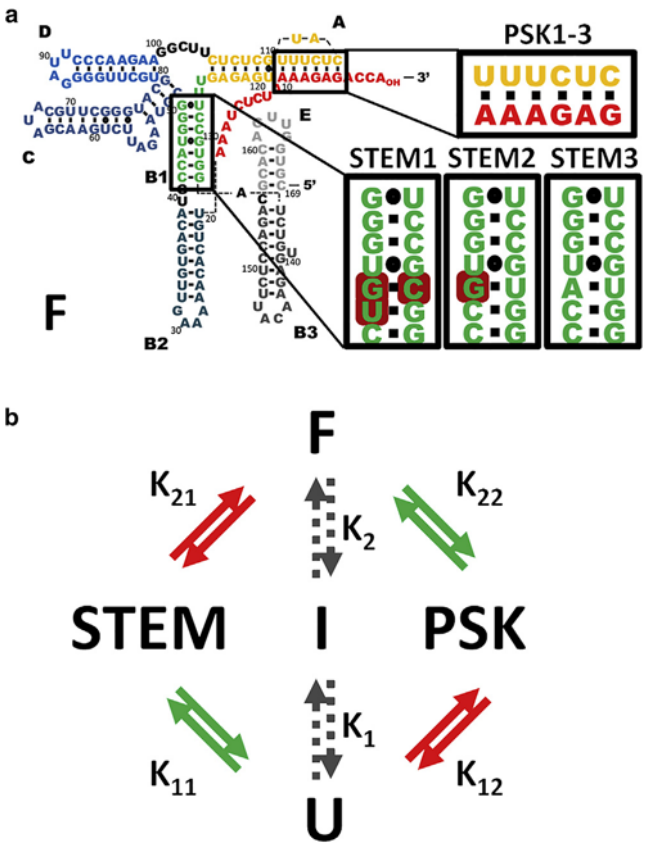

FIGURE 7 Effects of differences in the TLS domains of BMV1-3. (*a*) The 3′ 169-nt TLS domain is conserved across all three genomic RNAs. Compared to the above described BMV RNA3, the only changes observed in this region are located in the STEM domain. RNA2 contains one point mutation of A44 to G44. RNA1 contains three mutations of A44 to G44, C43 to G43, and U141 to C141. (*b*) Schematic of the stepwise folding pathway of TLS in BMV. The choice of the intermediate structure I depends on the equilibrium concentrations of STEM, PSK, and U conformations in solution. Such concentrations are given by the equilibrium constants K_{11} and K_{12} and therefore the energetics of formation of the STEM and PSK state. Destabilizing of the stem domain as observed in RNA1 and RNA2 thus shifts the equilibrium toward the PSK conformation and thus the PSK pathway. To see this figure in color, go online.

composition toward the PSK conformation, possibly making the PSK pathway a feasible folding alternative (Fig. 7 *b*).

STEM / F conformational switching as a regulatory mechanism

Previous structural models were based on the observation of the F conformer at high and the U conformer at low [Mg^{2+}]. The likely presence of both STEM and F conformers in equilibrium at pseudo-physiological conditions suggests an intriguing mechanism for regulation between RNA replication steps. It is worth noting that Felden et al. (23) have previously observed an alternate conformation in 3° structural modeling, and postulated that F could be in a dynamic equilibrium with a conformation they referred to as an open form. Interestingly, this open form agrees well with the STEM interaction in our folding model, in which the 3′ terminal pseudoknot is replaced with an alternate hairpin (*A*′, Figs. 1 and 2). In this conformation, the 3′-terminal-CCA_{OH} is presented in an alternate tertiary conformation and therefore inaccessible to pseudoknot specific host enzymes leading to loss of some biological functions, e.g., inactivity for amino acylation. Dreher and Hall (48) suggested that presence of both states allows shifting of the equilibrium toward the F conformer through selective stabilization upon binding to tRNA amino-acyl transferases. In fact, tRNA amino acylation has been observed to operate at stoichiometric efficiencies with BMV TLS as substrate, thus validating the conclusion of Dreher and Hall (48).

Going one step further, a complementary mechanism could be envisioned in which interaction with a host factor or viral protein shifts the equilibrium toward the STEM conformer. As the F conformer is required for recognition of replication-related host factors, this mechanism would be equivalent to turning the viral replication pathway off. In fact, a similar mechanism has been observed for members of the related *Ilarvirus* genus of the *Bromoviridae* family that share similarities to BMV, such as alfalfa mosaic virus (AMV) (49,50). The members of this genus require interaction of coat protein (CP) with a series of stem-loops near the 3′-end for infectivity (Fig. 8 *b*). This 3′-terminal domain is capable of folding into an alternative tRNA-like conformer that bears some resemblance to the TLS of BMV. As for BMV, the tRNA-like conformer is recognized by the viral

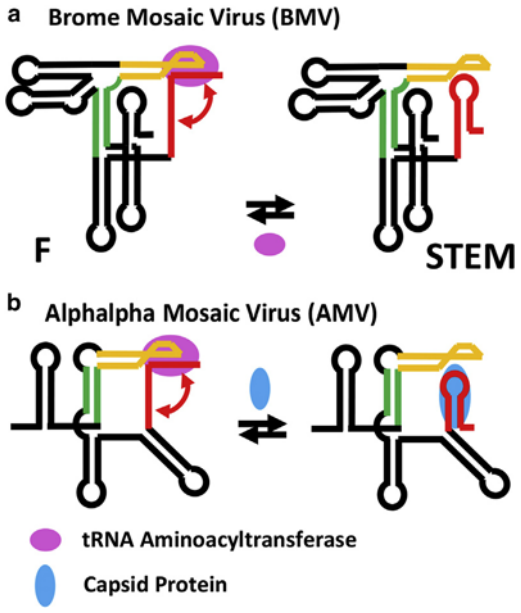

FIGURE 8 Structural similarities of BMV and AMV and known functional implications. (*a*) Structural switch in BMV. The TLS is charged with tyrosine by a tRNA amino-acyl transferase that recognizes the pseudoknot at the 3′ terminus. (*b*) Structural switch in AMV. The TLS in AMV responds to its capsid protein. Binding of the capsid protein to the 3′ hairpin in the alternate conformation induces assembly of the protein coat around the virus genome. To see this figure in color, go online.

replicase and thus required for minus-strand synthesis. Binding of the coat protein to the 3′-terminal hairpin results in disruption of the 3′ pseudoknot and thus inhibition of minus strand synthesis. Mg^{2+}, on the other hand, has been found to interfere with coat protein binding due to stabilization of the 3′ pseudoknot interaction (50). In agreement with such a mechanism, it appears as though BMV has evolved a complex scheme, resulting in delayed translation of its CP (51). Consequently, [CP] increases at later times during the viral replication cycle, thus presenting a viable factor for timed regulation of replication. Such a mechanism is quite analogous to the way that riboswitches turn off gene expression by binding to a downstream product, thus regulating expression in response to ligand binding.

In a similar fashion, the STEM # F conformational equilibrium could help in differential recognition of the individual RNA strands. In parallel to the predicted stabilities of the STEM domains (RNA1~RNA2 < RNA3), different biological properties are observed for RNA3 compared to RNA1 and RNA2. By means of specific example, RNA1 and RNA2 require 3′ amino acylation to support infectivity, but not RNA3. Contrariwise, RNA3 requires the TLS domain as a *cis*-regulating element for effective genome encapsidation, while it can be *cis*- or *trans*-regulating in RNA1 and RNA2 (8,27). Indeed, subtle differences in the STEM domains may therefore provide a novel pathway for differential performance between RNA3 and RNA1/RNA2 genomic strands, which is crucially needed for timing purposes during the viral replication cycle. However, clearly much more work remains to be done to establish the details of such a putative mechanism. In particular, time-resolved TIRF or confocal single-molecule studies into the folding rate constants for these TLS constructs would be extremely valuable, but to observe such slow conformational kinetics will require spatial tethering of these TLS domains on the coverslip surface. Work toward this goal is underway.

CONCLUSIONS

We have synthesized an RNA construct for single-molecule fluorescence experiments modeled after the TLS domain of Brome Mosaic Virus RNA3, which allows study of TLS folding dynamics via changes in FRET efficiencies caused by shifts in Cy3-Cy5 donor-acceptor distances. Burst fluorescence has been utilized to observe and interpret $[M^{n+}]$-induced folding of the TLS domain under freely diffusing conditions. Histograms of the E_{FRET} probabilities reveal three distinct conformations for TLS molecules in solution, in sharp contrast with the accepted two-state model (i.e., unfolded (U) # folded (F)) found in the literature. In particular, the fractional TLS populations exhibit a smooth progression with increasing $[M^{n+}]$ from unfolded (U) molecules, through a stable intermediate state (I), to formation of the fully folded conformation (F). As the folded F conformation is held in place by both stem and pseudoknot tertiary interactions, an equilibrium model has been developed that allows for sequential folding pathways through either STEM or PSK intermediates. Single molecule analysis of the fractional populations as a function of monovalent (Na^{+}) and divalent (Mg^{2+}) cations is found to be in excellent agreement with a three-state, sequential step model, dominated by formation of the STEM intermediate. This consecutive folding model is further validated by oligomer competition studies, which unambiguously identify the intermediate species as partial folding via the stem interaction.

Of particular importance is that the STEM intermediate represents a significant species under physiological relevant salt conditions. Detailed burst titration studies reveal the unfolded U conformational state to be of relatively little importance, but instead that both STEM and F conformations exist in a coupled equilibrium that depends sensitively on both $[Mg^{2+}]$ and $[Na^{+}]$. These studies suggest that differential interaction of these STEM and F conformers with host factors may present a possible mechanism for the regulation of the viral replication cycle. Alternatively stated, shifts in equilibrium properties between TLS domains may play a key role in biochemical recognition and differentiation of the RNA1 and RNA2 versus RNA3 genomic strands. However, further work will be required to establish or refute the validity of such mechanisms, as is being explored by kinetic studies of immobilized TLS constructs at the single molecule level.

AUTHOR CONTRIBUTIONS

M.V. and E.D.H. performed research; M.V., E.D.H., and D.J.N. analyzed data; M.V., E.D.H., and D.J.N. contributed analytic tools; M.V. and D.J.N. designed research; and M.V., E.D.H., and D.J.N. wrote the article.

ACKNOWLEDGMENTS

The authors thank Professor Robert T. Batey and Jake Polaski for help with the RNA synthesis.

Support for this work has been provided by the National Science Foundation (CHE1266416, PHY1125844), the National Institute of Standards and Technology, and the W. M. Keck Foundation initiative in RNA sciences at the University of Colorado, Boulder. The National Institutes of Health Molecular Biophysics Training Program (T32 GM-065103) provided partial support for E.D.H. GM-065103) provided partial support for E.D.H.

REFERENCES

1. Walter, N. G., S. A. Woodson, and R. T. Batey. 2009. Non-Protein Coding RNAs. Springer Series in Biophysics, Vol. 13. Springer, Berlin, Germany.
2. Jacobs, J., S. Glanz, . , U. Kück. 2010. RNA *trans*-splicing: identification of components of a putative chloroplast spliceosome. *Eur. J. Cell Biol.* 89:932–939.
3. Zong, X., V. Tripathi, and K. V. Prasanth. 2011. RNA splicing control: yet another gene regulatory role for long nuclear noncoding RNAs. *RNA Biol.* 8:968–977.
4. Collins, K. 2009. Forms and functions of telomerase RNA. *Springer Ser. Biophys.* 13:285–301.

5. Cech, T. R. 2009. Crawling out of the RNA world. *Cell.* 136:599–602.
6. Cech, T. R. 2013. How a chemist looks at RNA. *Angew. Chem. Int. Ed. Engl.* 52:75–78.
7. Serganov, A., and E. Nudler. 2013. A decade of riboswitches. *Cell.* 152:17–24.
8. Dreher, T. W. 2009. Role of tRNA-like structures in controlling plant virus replication. *Virus Res.* 139:217–229.
9. Dreher, T. W. 2010. Viral tRNAs and tRNA-like structures. *Wiley Interdiscip. Rev. RNA.* 1:402–414.
10. Gesteland, R. F., T. R. Cech, and J. F. Atkins. 1999. The nature of modern RNA suggests a prebiotic RNA. *In* The RNA World, 2nd Ed.. Cold Spring Harbor Monograph Series. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 37.
11. Batey, R. T., R. P. Rambo, and J. A. Doudna. 1999. Tertiary motifs in RNA structure and folding. *Angew. Chem. Int. Ed. Engl.* 38:2326–2343.
12. Tinoco, I., Jr., and C. Bustamante. 1999. How RNA folds. *J. Mol. Biol.* 293:271–281.
13. Draper, D. E. 2004. A guide to ions and RNA structure. *RNA.* 10: 335–343.
14. Souliere, M. F., A. Haller, T. Santner, and R. Micura. 2013. New insights into gene regulation – High resolution structures of cobalamin riboswitches. *Angew. Chem. Int. Ed. Engl.* 52:1874–1877.
15. Voet, D., and J. G. Voet. 1995. Biochemistry, 2nd Ed. John Wiley, New York.
16. Grigg, J. C., and A. Ke. 2013. One platform, five brands: how nature cuts the cost on riboswitches. *J. Mol. Biol.* 425:1593–1595.
17. Soulière, M. F., A. Haller, . . , R. Micura. 2013. New insights into gene regulation–high-resolution structures of cobalamin riboswitches. *Angew. Chem. Int. Ed. Engl.* 52:1874–1877.
18. Simon, A. E., and L. Gehrke. 2009. RNA conformational changes in the lifecycles of RNA viruses, viroids, and virus-associated RNAs. *Biochim. Biophys. Acta.* 1789:571–583.
19. Dimmock, N. J., A. J. Easton, and K. Leppard. 2001. Introduction to Modern Virology. Blackwell Publishing, Oxford, UK.
20. Dreher, T. W. 1999. Functions of the 3'-untranslated regions of positive strand RNA viral genomes. *Annu. Rev. Phytopathol.* 37:151–174.
21. Mans, R. M. W., C. W. A. Pleij, and L. Bosch. 1991. tRNA-like structures. Structure, function and evolutionary significance. *Eur. J. Biochem.* 201:303–324.
22. Rao, A. L. N., and C. Cheng Kao. 2015. The brome mosaic virus 3' untranslated sequence regulates RNA replication, recombination, and virion assembly. *Virus Res.* 206:46–52.
23. Felden, B., C. Florentz, . . , E. Westhof. 1994. Solution structure of the 3′-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. *J. Mol. Biol.* 235:508–531.
24. Kao, C. C., and K. Sivakumaran. 2000. Brome Mosaic Virus, good for an RNA virologist's basic needs. *Mol. Plant Pathol.* 1:91–97.
25. Young, M., D. Willits, . . , T. Douglas. 2008. Plant viruses as biotemplates for materials and their use in nanotechnology. *Annu. Rev. Phytopathol.* 46:361–384.
26. Ahlquist, P., R. Dasgupta, and P. Kaesberg. 1981. Near identity of 3-RNA secondary structure in bromoviruses and cucumber mosaic virus. *Cell.* 23:183–189.
27. Annamalai, P., and A. L. N. Rao. 2007. In vivo packaging of Brome Mosaic Virus RNA3, but not RNAs 1 and 2, is dependent on a *cis*-acting 3′ tRNA-like structure. *J. Virol.* 81:173–181.
28. Hammond, J. A., R. P. Rambo, . . , J. S. Kieft. 2009. Comparison and functional implications of the 3D architectures of viral tRNA-like structures. *RNA.* 15:294–307.
29. Rambo, R. P., and J. A. Tainer. 2010. Improving small-angle x-ray scattering data for structural analyses of the RNA world. *RNA.* 16:638–646.
30. Rietveld, K., C. W. A. Pleij, and L. Bosch. 1983. Three-dimensional models of the tRNA-like 3′ termini of some plant viral RNAs. *EMBO J.* 2:1079–1085.
31. Dahan, M., A. A. Deniz, . . , S. Weiss. 1999. Ratiometric measurement and identification of single diffusing molecules. *Chem. Phys.* 247: 85–106.
32. Deniz, A. A., T. A. Laurence, . . , S. Weiss. 2001. Ratiometric single-molecule studies of freely diffusing biomolecules. *Annu. Rev. Phys. Chem.* 52:233–253.
33. Deniz, A. A., M. Dahan, . . , P. G. Schultz. 1999. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. USA.* 96:3670–3675.
34. Grunwell, J. R., J. L. Glass, . . , P. G. Schultz. 2001. Monitoring the conformational fluctuations of DNA hairpins using single-pair fluorescence resonance energy transfer. *J. Am. Chem. Soc.* 123:4295–4303.
35. Bokinsky, G., and X. Zhuang. 2005. Single-molecule RNA folding. *Acc. Chem. Res.* 38:566–573.
36. Rothwell, P. J., S. Berger, . . , C. A. M. Seidel. 2003. Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase:primer/template complexes. *Proc. Natl. Acad. Sci. USA.* 100:1655–1660.
37. Chapman, M. R., and C. C. Kao. 1999. A minimal RNA promoter for minus-strand RNA synthesis by the Brome Mosaic Virus polymerase complex. *J. Mol. Biol.* 286:709–720.
38. Hermanson, G. 2013. Bioconjugate Techniques, 3rd Ed. Academic Press, Waltham, MA.
39. Fiore, J. L., E. D. Holmstrom, and D. J. Nesbitt. 2012. Entropic origin of Mg^{2+}-facilitated RNA folding. *Proc. Natl. Acad. Sci. USA.* 109:2902–2907.
40. Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* 94:1826–1835.
41. Fiore, J. L., J. H. Hodak, . . , D. J. Nesbitt. 2008. Monovalent and divalent promoted GAAA tetraloop-receptor tertiary interactions from freely diffusing single-molecule studies. *Biophys. J.* 95:3892–3905.
42. Fiegland, L. R., A. D. Garst, . . , D. J. Nesbitt. 2012. Single-molecule studies of the lysine riboswitch reveal effector-dependent conformational dynamics of the aptamer domain. *Biochemistry.* 51:9223–9233.
43. Ha, T., X. Zhuang, . . , S. Chu. 1999. Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. USA.* 96:9077–9082.
44. Leipply, D., D. Lambert, and D. E. Draper. 2009. Ion-RNA interactions thermodynamic analysis of the effects of mono- and divalent ions on RNA conformational equilibria. *Methods Enzymol.* 469:433–463.
45. Fiore, J. L., E. D. Holmstrom, . . , D. J. Nesbitt. 2012. The role of counterion valence and size in GAAA tetraloop-receptor docking/undocking kinetics. *J. Mol. Biol.* 423:198–216.
46. Turner, D. H., N. Sugimoto, and S. M. Freier. 1988. RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.* 17:167–192.
47. Bokinsky, G., D. Rueda, . . , X. Zhuang. 2003. Single-molecule transition-state analysis of RNA folding. *Proc. Natl. Acad. Sci. USA.* 100:9302–9307.
48. Dreher, T. W., and T. C. Hall. 1988. Mutational analysis of the tRNA mimicry of brome mosaic virus RNA. Sequence and structural requirements for aminoacylation and 3′-adenylation. *J. Mol. Biol.* 201:41–55.
49. Chen, S.-C., and R. C. L. Olsthoorn. 2010. In vitro and in vivo studies of the RNA conformational switch in Alfalfa mosaic virus. *J. Virol.* 84:1423–1429.
50. Olsthoorn, R. C. L., S. Mertens, . . , J. F. Bol. 1999. A conformational switch at the 3′ end of a plant virus RNA regulates viral replication. *EMBO J.* 18:4856–4864.
51. Kao, C. C., P. Ni, . . , B. Dragnea. 2011. The coat protein leads the way: an update on basic and applied studies with the Brome Mosaic Virus coat protein. *Mol. Plant Pathol.* 12:403–412.