HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway

Hansini Mundigala1, Jonathan B. Michaux1, Andrew L. Feig1,*, Eric Ennifar2,* and David Rueda1,3,4,*

1Department of Chemistry, Wayne State University, Detroit, MI 48236, USA, 2Architecture et Réactivité de l’ARN, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire du CNRS, F-67084 Strasbourg, France, 3Department of Medicine, Section of Virology, Imperial College, London W12 0NN, UK and 4Single Molecule Imaging Group, MRC Clinical Sciences Center, Imperial College, London W12 0NN, UK

Received November 30, 2013; Revised March 24, 2014; Accepted April 8, 2014

ABSTRACT

The HIV-1 dimerization initiation sequence (DIS) is a conserved palindrome in the apical loop of a conserved hairpin motif in the 5′-untranslated region of its RNA genome. DIS hairpin plays an important role in genome dimerization by forming a ‘kissing complex’ between two complementary hairpins. Understanding the kinetics of this interaction is key to exploiting DIS as a possible human immunodeficiency virus (HIV) drug target. Here, we present a single-molecule Förster resonance energy transfer (smFRET) study of the dimerization reaction kinetics. Our data show the real-time formation and dissociation dynamics of individual kissing complexes, as well as the formation of the mature extended duplex complex that is ultimately required for virion packaging. Interestingly, the single-molecule trajectories reveal the presence of a previously unobserved bent intermediate required for extended duplex formation. The universally conserved A272 is essential for the formation of this intermediate, which is stabilized by Mg2+, but not by K+ cations. We propose a 3D model of a possible bent intermediate and a minimal dimerization pathway consisting of three steps with two obligatory intermediates (kissing complex and bent intermediate) and driven by Mg2+ ions.

INTRODUCTION

Human immunodeficiency virus (HIV), a retrovirus, attacks the human immune system, which can result in acquired immune deficiency syndrome (AIDS) (1–3). HIV is the leading cause of death in Africa and the fourth leading cause of death worldwide (4,5). Current therapies against HIV target mainly two viral enzymes: reverse transcriptase (6) and protease (7). Due to the rapid evolution of strains resistant to enzymatic inhibitors, new targets must be identified.

HIV contains two similar copies of its genomic RNA (gRNA), which share numerous intermolecular interactions (8). The most prominent one is the dimer linkage structure (DLS) in the 5′-untranslated region (UTR) of the viral genome (9,10), which has been shown to control translation, RNA packaging and recombination during proviral DNA synthesis. It has also been shown that multiple structural transitions in the 5′ UTR can regulate gRNA packaging (11,12).

Within the DLS, a highly conserved, nine-nucleotide apical loop, including a six-nucleotide, palindromic, dimerization initiation sequence (DIS), has been shown to be important in the dimerization process. The palindromic sequence is flanked by three conserved purines (A272, A/G273 and A280) that are essential for the dimer stability (13–15). The interaction between the DLS of the two gRNA is initiated by formation of a kissing loop at the DIS (16,17). Mutations or alterations of the DIS prevents RNA dimerization and severely reduces the viral infectivity (18–21). It has been shown that the 35 nucleotide (nt) DLS with the lower stem bulge is required for the two-step dimerization in presence of NCp7 whereas the 23 nt construct can achieve the two-step dimerization in absence of this protein (18–21). Experiments on synthetic RNA fragments have shown that the initial kissing-loop dimer is subsequently stabilized by extension of intermolecular Watson–Crick base pairs as an extended duplex (ED) (16,25). This kissing loop to ED isomerization is strongly facilitated in vitro by incubation at high temperature (55°C), or at physiological temperature by the nucleocapsid protein (NC), a small, basic protein with two zinc-finger domains (22,23,25–29). Structural insights of the DIS kissing loop and ED forms have been provided by X-ray crystallography (30–32) and nuclear magnetic res-
on 06 February 2018

Figure 1. smFRET detection of the bimolecular HIV-1 RNA dimerization. (a) Secondary structure of the RNA strands used in this study with the fluorophores. A single base-flip mutation in the center base pair of the loop is used to prevent homodimerization and promote dimerization of Cy3–Cy5 labeled hairpins. (b) Schematic diagram of the single molecule experiments. A Cy3 labeled DIS-1 hairpin is surface-immobilized onto the quartz slide via a biotin-streptavidin bridge. Cy5 labeled DIS-2 hairpin is introduced into the slide and allowed to interact freely with the immobilized DIS-1 under near-physiological conditions (20 mM TRIS pH 7.4, 5 mM Mg\(^{2+}\) and 150 mM KCl). The fluorophores are excited via a prism-based total internal reflection microscope. Fluorescence is collected through the objective, separated with appropriate dichroic mirrors and monitored with a CCD camera. The dimerization pathway consists of three possible FRET states: free DIS-1 (0 FRET), the kissing complex (0.4 FRET) and the extended duplex (1.0 FRET). (c) Resulting FRET time trajectory shows the formation (FRET ~ 0.4) and dissociation (FRET ~ 0) of the kissing-loop complex. Dwell times in each state (\(\tau_{on}\) and \(\tau_{off}\)) are used to build the dwell time distribution (see Figure 2).

MATERIALS AND METHODS

RNA purification and labeling

RNA samples were purchased from the Keck Foundation Resource Laboratory at the Yale University School of Medicine with a 5′-Cy3 and a 3′-biotin (DIS-1, Figure 1a) or with a 3′-C7 amino linker (DIS-2, Figure 1a). An intermolecular GC base pair in the loop was flipped to prevent homodimer formation while maintaining the KC stability (Figure 1a and Supplementary Materials). RNAs were deprotected, purified and labeled as previously described (40,41). Briefly, deprotected RNAs were purified by denaturing gel electrophoresis (20% wt/vol polyacrylamide and 8 M urea) and diffusion elution against elution buffer (0.5 M NH\(_4\)OAc and 0.1 mM EDTA) overnight at 4°C, followed by chloroform extraction, ethanol precipitation and C8 reverse-phase High Performance Liquid Chromatography (HPLC). The C7 amino linker in DIS2 was labeled with Cy5 (GE Healthcare) in labeling buffer (100 mM Na\(_2\)CO\(_3\), pH 8.5) overnight at room temperature. The labeled RNA was further purified by ethanol precipitation and reverse-phase HPLC using a linear gradient of acetonitrile in triethylammonium acetate as mobile phase, as described (40). RNA concentrations were determined by UV-Vis absorbance at 260 nm and absorbance at wavelengths of 550 nm and 650 nm were obtained to quantify Cy3 and Cy5 dye incorporation.

Single-molecule FRET

Single molecule experiments were performed as previously described (42,43). Briefly, RNA strands DIS-1 and DIS-2 were diluted to 25 pM and 30 nM, respectively, in standard buffer (20 mM Tris-HCl pH 7.5, 150 mM KCl and 5 mM MgCl\(_2\) in saturating trolox). DIS-1 was heated at 90°C for 45 s before flash cooling on ice to prevent homodimerization. DIS-1 RNA was surface-immobilized onto streptavidin-coated quartz slide via a biotin-streptavidin bridge (Figure 1b) to generate a surface density of ~0.1 molecules/μm\(^2\). DIS-2 was manually flowed into the slide microchannel in standard buffer with an oxygen-scavenging system (5 mM 3,4-dihydroxybenzoic acid and 0.1 μM protocatechuate 3,4-dioxigenase) to minimize photobleaching. Cy3 was excited in a home-built total internal reflection microscope with a laser (532 nm, 3 mW, Spectra-Physics, Excelsior). Donor and acceptor emissions were separated using appropriate dichroic mirrors (610DCXR, Chroma) and detected as two side-by-side images on a back-illuminated electron-multiplied CCD camera (I-Xon, Andor). Measurements were obtained under variable [Mg\(^{2+}\)] and [K\(^{+}\)] (0.001–20 mM and 0–1 M, respectively) at room temperature. Dynamic FRET traces were rarely observed in the time duration of experiments (15–20 min) with a time resolution of 33 ms. In order to maximize the time duration of experiments, 1000 ms time resolution was used and time trajectories were recorded with one frame per second time resolution for up to 45–60 min (Figure 1c). Trajectories were time-binned to construct FRET histograms, and dwell times calculated for each dimerization event to determine the rate constants, as described (45). FRET cutoff values of 0.25 and 0.75 were used to distinguish between monomer and KC and ED. Dwell time histograms were fit to either single or double exponentials to determine \(k_{on}\) and \(k_{off}\). Metal ion (Mg\(^{2+}\) and K\(^{+}\)) titrations were fit to the modified Langmuir...
Equation (1) to obtain the binding constants \( (K_{1/2}) \):

\[
f(x) = f_0 + (f_{\text{max}} - f_0) \frac{x}{K_{1/2} + x},
\]

where \( f_0 \) and \( f_{\text{max}} \) are the initial and saturating populations, respectively, and \( x \) is the concentration of Mg\(^{2+} \) or K\(^+ \) ions.

**Model building**

Modeling was carried out starting from the HIV TAR RNA/SELEX RNA kissing-loop complex (PDB ID 2RN1) (44) using the program Coot (46). Residues of the stem and the loop were mutated to match the DIS sequence and each stem length was extended to seven base pairs. Stem/loop junctions were opened in order to insert flanking adenines 5’ and 3’ of the loop, without changing the angle between each stem. Adenines 280 were modeled stacked inside the helix whereas adenines 272 and 273 were placed outside the helix and stacked on each other, similarly to their respective conformation observed in the HIV DIS kissing-loop complex crystal structure (PDB ID 1XPF) (32). After model building, several cycles of geometrical regularization were carried out with Coot.

**RESULTS**

smFRET reveals that kissing complex formation and dissociation is an extraordinarily slow process

To dissect the in vitro dimerization reaction of DIS RNA (Figure 1a), we have used a smFRET assay previously developed to monitor the KC formation of RNA hairpins in general (45,47,48). The FRET donor-labeled hairpin (DIS1) is surface-immobilized via a biotin-streptavidin bridge while the FRET acceptor-labeled hairpin (DIS2) diffuses in standard buffer under near-physiological conditions (Figure 1b). A typical single molecule FRET time trajectory (Figure 1c) reveals the presence of random excursions between two distinct states at \( \sim 0 \) and \( \sim 0.4 \) FRET, corresponding to DIS1 monomer and the DIS1:DIS2 KC, respectively. Using the latter FRET value and typical approximations (\( k^2 = 2/3 \) and \( R_0 = 60 \) Å for Cy3–Cy5) (49), we estimate the distance between the two fluorophores to be \( \sim 64 \) Å, in excellent agreement with the KC crystal structure (31). The time trajectories reveal that the hairpins can undergo many cycles of association and dissociation without progressing to ED formation and that the KC can be very long-lived (greater than 10 min). A dwell-time analysis in the monomer (\( \tau_{\text{off}} \)) and KC (\( \tau_{\text{on}} \)) states yields the pseudo-first-order binding rate constant \( k'_{\text{on}} \) and dissociation (\( k_{\text{off}} \)) rate constants, respectively (Figure 2a). The resulting histograms confirm the slow association and dissociation kinetics, as well as heterogeneous dissociation kinetics (double exponential), indicating that two KC populations with different stabilities (i.e. loop–loop interactions) may be formed. As expected for a binary reaction, \( k_{\text{off}} \) is independent of the RNA concentration, while \( k'_{\text{on}} \) increases linearly with [DIS2] (Figure 2b). A linear fit to the latter yields the second-order binding rate constant \( k_{\text{on}} = 10^5 \text{M}^{-1}\text{s}^{-1} \). This extremely slow formation rate constant may be the result of structural rearrangement of the hairpin prior to forming the KC, which makes most diffusion-controlled collisions between the two hairpins unsuccessful.

To further investigate the origin of the two KC populations, we conducted metal ion titrations.

Mg\(^{2+} \) ions stabilize the kissing complex

To test whether the observed heterogeneity results from different metal ion interactions, we measured \( k'_{\text{on}} \) and \( k_{\text{off}} \) as a function of [Mg\(^{2+} \)] and [K\(^+ \)] (Figure 3). The observed \( k'_{\text{on}} \) is independent of both [Mg\(^{2+} \)] and [K\(^+ \)] (Figure 3, top). Interestingly, the long-lived KC population with a slow dissociation constant (\( k_{\text{off},2} \)) is only observed at \([\text{Mg}^{2+}] \geq 0.5 \text{ mM}\), whereas the short-lived KC population (\( k_{\text{off},1} \)) is observed across the entire \([\text{Mg}^{2+}] \) range (Figure 3a, middle). The data show that the magnitude of both \( k_{\text{off},1} \) and \( k_{\text{off},2} \) remains invariant with [Mg\(^{2+} \)], but the fraction of the short-lived KC population (\( f_j \)) decreases sharply near 0.5 mM [Mg\(^{2+} \)] in favor of the long-lived population (Figure 3a, bottom). This result implies that the slow population arises from binding of a specific magnesium ion to the hairpin with a dissociation constant of 0.5 mM and stabilizing the KC by \( \sim 2 \) kcal mol\(^{-1} \). This result is in agreement with the presence of a Mg\(^{2+} \) ion in the crystal structure of the DIS KC and with prior biochemical and biophysical bulk experiments (50,51).

To test for the specificity of this magnesium ion interaction, we titrated potassium ions in a background of 5 mM Mg\(^{2+} \). The data show that, above 300 mM, K\(^+ \) ions can partially recover the fast dissociating population (\( f_j \)) indicating that monovalent ions can compete against the divalent ion for the binding site. However, the large amount of K\(^+ \)
ions required for this competition supports the notion that a tightly and specifically bound Mg$^{2+}$ ion is primarily responsible for the observed high stability of the KC.

**smFRET reveals an obligatory bent intermediate**

Our labeling strategy enables us to study the dynamic behavior of the KC as well as its progression toward the extended RNA duplex (ED, Figure 1). With this labeling scheme, ED is expected to result in a high static FRET state, while the KC is expected to yield a mid FRET state. A smFRET time trajectory (Figure 4a, left) shows the progression to the ED conformation (static FRET = 1.0) following multiple associations and dissociations of monomer hairpins to KC. Under standard conditions (5 mM Mg$^{2+}$, 150 mM K$^+$, 20 mM TRIS pH 7.5, 22°C), only 13% of trajectories reach the ED state. However, under these conditions, we also observe 32% of trajectories exhibiting dynamic excursions between FRET 1.0 and 0.4 (Figure 4a, right). Control experiments confirm that this behavior is never observed with pre-annealed ED complexes, ruling out possible photophysical artifacts caused by the local environment in the ED configuration. Therefore, this dynamic population is likely to represent a different folded state that brings the fluorophores in close proximity, such as a bent KC (similar to the TAR–TAR* complex (52–54)) or a cruciform intermediate initiated by fraying of the hairpins’ stems (55).

The fraction of trajectories exhibiting this intermediate conformation increases with Mg$^{2+}$ ion concentration, indicating that Mg$^{2+}$ ions stabilize this population (Figure 4b, left). Measuring the intermediate docking and undocking rate constants as a function of RNA concentration (Figure 4c and Supplementary Figure S6) shows that docking and
Magnesium ions are required to form the bent intermediate

To assess the effect of monovalent and divalent ions on the folding pathway, we determined the fraction of each observed state (KC, intermediate bent conformation and extended) with varying [K+] and [Mg2+] in the presence of saturating concentrations of free DIS2. The Mg2+ titration was performed at 150 mM background K+ ions, while the K+ titration was performed in a 5 mM background of Mg2+ ions. At low [Mg2+], ∼75% of molecules adopt the KC conformation, and only 10–20% form the intermediate or the ED conformation. Increasing the [Mg2+] to 50 mM (Figure 4b, left panel) shifts the KC population to the intermediate state, making both fractions equally populated (40%). This data suggest that Mg2+ ions promote conversion of KC into the bent intermediate. At high monovalent concentrations (>500 mM), the amount of bent intermediate is highly reduced (∼15%), confirming that Mg2+ is required for the transition into the intermediate form.

A stem mutant rules out a possible cruciform intermediate

To characterize the structure of the intermediate and to distinguish between the possible conformations that bring the fluorophores in close proximity, we designed a hairpin mutant where the three terminal base pairs were flipped (Figure 5a), which prevents the formation of the ED through a cruciform (55) or any other intermediate. Experiments with this mutant reveal that 35% of molecules still exhibit dynamic excursions between FRET 1.0 and FRET 0.4. This result clearly rules out the formation of a cruciform intermediate in our experimental conditions and supports the bent KC conformation (Figure 5a). Additional experiments with the mutant also reveal that the excursion times in the FRET 1.0 state are reduced relative to the wild type (Figure 4 and Supplementary Table S1), suggesting that long range interactions between the complementary stem regions of the RNA hairpins in KC may contribute to the stability of the bent kissing structure (Figure 5). To further test this model, we introduced both donor and acceptor fluorophores on a single hairpin (DIS-2) and observed the ED formation (Supplementary Figure S7). In this scheme, the model predicts that both the KC and the bent complex are characterized by high FRET values (the stem remains base paired), whereas the ED should yield intermediate FRET. The data support the kinetic model and shows that the stem regions of the hairpin RNA remain base paired until the irreversible formation of the ED.

In summary, these data suggest that the dynamic intermediate transiently adopts a bent KC conformation (FRET 1.0) in equilibrium with the coaxially stacked conformation (FRET 0.4, Figure 1b).

3D model of a DIS kissing-loop bent intermediate

 Modeling of a possible bent DIS KC was carried out starting from the kinked TAR–TAR* kissing-loop NMR structure and the coaxially stacked DIS kissing-loop complex crystal structure. Structure modeling showed that only limited accommodations are required to induce a significant bending of the coaxially stacked KC. In particular, flanking adenines 272 and 2735‘ of the self-complementary sequence can maintain an extra-helical conformation in the bent KC similar to the one observed in the coaxial one (Figure 6).

A272 is essential for stability and bending of the KC

It has been established that the stability of the KC is strongly dependent on the three highly conserved purines flanking the self-complementary sequence (13,14). In particular, it was suggested that A272 has a large impact on the dynamics and local conformational changes of the KC (14,24). We investigated the effect of A272 by mutating this base to a cytidine (A272C, Figure 5b). Under standard buffer conditions, the A272C mutant exhibits no long-lived KC population (Supplementary Table S1), similar to the wild-type hairpins in the absence of Mg2+ ions. This result is consistent with the binding of a Mg2+ ion in the vicinity of A272, thus stabilizing the flipped out conformation of the base. This, in turn, can stabilize the KC through inter-helical adenine stacking interactions as observed in crystal structures (31).

Furthermore, the introduction of A272C mutation to one of the RNA hairpins significantly decreases (<10%) the population of molecules in the bent intermediate confor-
dimerization. OursmFRETkineticdatarevealthatthefor-

minimal RNA sequence responsible for the viral genome
dimerization mechanism of HIV-1 in real time using the

tier for the conversion of KC to the ED (48). RNA hairpins have shown that there is a large kinetic bar-
exact mechanism remained unclear. Studies using model

iesconcludedthatisomerizationproceedswithoutcomplete

disruption of the loop–loop helix. However, details of the

Figure 6. 3D architecture of HIV-1 DIS RNA dimers along the isomer-

ization pathway. The two strands are shown in orange and blue. (a) The

DIS kissing complex as observed in crystal structures (PDB 1XPF). (b)

Molecular model of a bent kissing complex based on a TAR complex (44).

(c) Extended duplex as observed in crystal structures (PDB 462D). Hex-

ahydrated magnesium ions observed in X-ray structures are shown as red

spheres.

Figure 7. Minimal kinetic model for the HIV-1 RNA dimerization with

the observed bent dimer. First the KC forms, followed by the formation of

a dynamic bent intermediate, which eventually leads to the formation of

a stable extended duplex. Magnesium ions (green spheres) favor the bent

KC intermediate formation.

mation. This asserts the fact that the metastable KC is an

obligatory step that is required for the transition into the

bent KC.

DISCUSSION

RNA KCs/loop–loop interactions play a major role in mul-
tistep RNA folding pathways. Kinetics and thermodynam-
ics of RNA loop–loop interactions have been previously
evaluated using multiple techniques such as surface plas-
mon resonance (SPR), electrospray ionization mass spec-
trometry (ESI-MS), ITC, NMR, bulk fluorescence mea-
surements, electrospray ionization-Fourier transform mass
spectrometry and UV melting (23,24,48,55–58). These stud-
ies concluded that isomerization proceeds without complete
disruption of the loop–loop helix. However, details of the
exact mechanism remained unclear. Studies using model
RNA hairpins have shown that there is a large kinetic bar-
rier for the conversion of KC to the ED (48).

In this study, we have used smFRET to characterize the
dimerization mechanism of HIV-1 in real time using the
minimal RNA sequence responsible for the viral genome
dimerization. Our smFRET kinetic data reveal that the for-
mation and dissociation of HIV-1 KC is an extraordinarily slow process under near-physiological conditions in vitro.

We observed multiple association and dissociation steps from the monomer to the KC dimer. The dissociation kinetic
is highly dependent on salt conditions. In the presence of Mg2+, highly stable KC dimers were observed, which
agrees with previous reports showing that metastable kiss-

ing dimers are formed in the presence of divalent metal ions
(23). It has been also shown that, in the presence of Mg2+, conversion of kissing dimer to the ED requires nucleocapsid protein (23). Our results clearly show that the conversion of KC to ED can be achieved slowly in the absence of nucleo-
capsid protein after multiple dissociation and re-association
steps. Our results with the A272C mutation shows that,
in the absence of A272, kissing dimers are unable to gain
high stability even in the presence of high concentrations
of [Mg2+]. This confirms the essential role played by this
universally conserved adenine in the formation of a sta-
ble kissing-loop complex (15,24,31). We propose that ade-
nine base flipping and stabilization of flipped adenines upon
Mg2+ binding might be the cause of the observed high sta-

bility of the kissing dimer in the presence of Mg2+. In sup-
port of this hypothesis, it has been shown that A272 can be
protonated and this might promote the loop dynamics and
the conversion to ED (24).

The smFRET trajectories also reveal the presence of an
intermediate in the transition from KC to the extended
RNA duplex. Experiments performed with inverted DIS
stem sequences clearly eliminate the possibility of cruciform
intermediates as observed with monomeric DIS hairpin mu-

nants (55). Based on the smFRET data, we propose that the

folding intermediate corresponds to a bent KC (Figure
6 and 7), similar to a TAR complex structure (44) or ColEl
plasmid specific RNA I and RNA II transcripts (59,60). A
similar concept of a bent KC retaining the WC base pairing
at the loop interaction and the C2 symmetry of the loop–
loop interface that facilitates the duplex formation via KC
has been suggested previously based on NMR data (61). A
similarly bent KC transition state has also been proposed in
the NCP7-chaperoned dimerization pathway (23). Alterna-

tively, the bent intermediate may resemble a structure sug-
gested in a recent NMR study, where the DIS hairpin base
pairing remains intact and inter-stem interactions are facil-
itated as a result of KC bending, which brings the two stems
in close proximity (62). In addition, there is in vivo evidence
for an intermediate HIV-1 gRNA dimer on the path from
immature gRNA dimer to mature gRNA dimer inside the
HIV-1 particle (63). We built a molecular model of a DIS
KC intermediate based on the TAR complex (44) by keep-
ing the angle between both hairpins constant (Figure 6).
In this model, A272 and A273 are in a flipped out configura-
tion and stacked onto each other as observed in KC crys-
tal structures, showing the feasibility of bending the DIS
KC with perfectly coaxially stacked hairpins thanks to the
plasticity provided by unpaired adenines. This is in line with
previous studies highlighting the dynamics of these purines
within the KC (24,32,64). This molecular model yields an
inter-stem distance of ~49 Å. However, based on the mea-
sured FRET efficiency (0.9–1.0), we estimate the distance
between the ends of the two hairpin stems to be ≤41 Å (as-
summing κ² = 24). Therefore, we propose that the observed bent KC adopts a conformation in between the model in Figure 6 and the completely bent KC conformation proposed by Dethoff et al. (62).

Monovalent and divalent ion titrations clearly show that Mg²⁺ ions are required for the formation of the proposed bent KC. Our data also highlight the important role of the universally conserved adenine 272, S′ of the loop. Experiments performed with the DIS A272C mutant show that the KC formation is strongly hindered, thus transition to bent KC or the ED cannot be readily achieved even at high Mg²⁺ concentrations.

CONCLUSIONS

The ability to characterize the folding pathway of HIV-1 RNA in vitro is critical to understanding the viral RNA dimerization during viral assembly. We have demonstrated a single-molecule fluorescence resonance energy transfer assay to monitor the dimerization of minimal HIV-1 RNA sequence containing DIS. Our smFRET data revealed that the bimolecular association rate constant of the two hairpin RNAs is 1.3 × 10⁴ M⁻¹ s⁻¹. This formation rate is independent of the Mg²⁺ and increases linearly with RNA concentration, which confirms that the formation is a diffusion-controlled reaction. The important role of Mg²⁺ for the DIS dimerization has been established by various bulk analysis methods (22,23,50,65). Our study supports the hypothesis that, in solution, magnesium binds the DIS with an equilibrium dissociation constant near 5 mM. Magnesium binding stabilizes the kissing interaction and the KC dissociation rates significantly increase (1.5 ± 0.3 min⁻¹) in the absence of magnesium. We observe magnesium bound metastable KC population with extreme slow dissociation rate (0.12 ± 0.02 min⁻¹). Our smFRET analysis reveals that HIV-1 RNA dimerization occurs through a three-step folding pathway in which the RNA KC shifts to a bent kissing conformation that leads to the formation of the extended RNA duplex via interaction through stems. We propose a 3D model of a possible bent DIS KC intermediate as expected from smFRET data. Our data also confirm that docking of the KC to form a bent conformation is independent of Mg²⁺ and RNA concentrations, although it requires the presence of Mg²⁺, indicating that bending is a unimolecular process. During maturation of the viral particle, the nucleocapsid protein (NCP7) chaperones the dimerization pathway. Thus, ongoing experiments are currently focused at studying the role of NCP7 in the dimerization pathway at single-molecule level. The mechanistic insights gained from these experiments represent significant progress toward understanding the HIV-1 dimerization mechanism and might help the rational development of new ligands targeting the HIV-1 DIS RNA (66).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

We would like to thank P. Dumas and C. Chow for helpful discussions and for commenting on the manuscript.

FUNDING

National Institute of Health [R01 GM085116]; National Science Foundation CAREER Award [MCB-0747285 to D.R.]. Agence Nationale pour la Recherche (ANR-12-BS07-0007-03 “ClickEnARN” to E.E.). Source of open access funding: Imperial College London.

Conflict of interest statement. None declared.

REFERENCES

1. Bender, W. and Davidson, N. (1976) Mapping of poly(A) sequences in oncornavirus RNA molecules. Cell, 7, 595–607.
2. Goto, T., Nakai, M. and Ikuta, K. (1998) The life-cycle of human immunodeficiency virus type 1. Micron, 29, 123–138.
3. Craveiro, M., Clerc, L., Sitbon, M. and Taylor, N. (2013) Metabolic pathways as regulators of HIV infection. Curr. Opin. HIV AIDS, 8, 182–189.
4. Mokili, J. and Korber, B. (2005) The spread of HIV in Africa. J. Neurovirol., 11(Suppl. 1), 65–77.
5. Cherutich, P., Bunnell, R. and Mermin, J. (2013) HIV testing: current practice and future directions. Curr. HIV/AIDS Rep., 10, 134–141.
6. Safranios, S.G., Marchand, B., Das, K., Himmel, D.M., Parniak, M.A., Hughes, S.H. and Arnold, E. (2009) Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. J. Mol. Biol., 385, 693–713.
7. Huff, J.R. (1991) HIV protease: a novel chemotherapeutic target for AIDS. J. Med. Chem., 34, 2305–2314.
8. Cann, A.J. and Karn, J. (1989) Molecular biology of HIV: new insights into the virus life-cycle. AIDS, 3(Suppl. 1), S19–S34.
9. Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B. and Ehresmann, C. (1994) Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc. Natl. Acad. Sci. U.S.A., 91, 4945–4949.
10. Laugna, M. and Jette, L. (1994) A 19-nucleotide sequence upstream of the 5′ major splice donor site is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. Biochemistry, 33, 13464–13474.
11. Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaiavel, G., Jones, S., Hiremath, A. et al. (2011) NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. Science, 334, 242–245.
12. Kenyon, J.C., Prestwood, L.J., Le Grice, S.F. and Lemon, A.M. (2013) In-gel probing of individual RNA conformers within a mixed population reveals a dimerization structural switch in the HIV-1 leader. Nucleic Acids Res., 41, e174.
13. Clever, J.L., Wong, M.L. and Parslow, T.G. (1996) Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. J. Virol., 70, 5902–5908.
14. Paillart, J.C., Westhof, E., Ehresmann, C., Ehresmann, B. and Marquet, R. (1997) Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA. J. Mol. Biol., 270, 36–49.
15. Lodmell, J.S., Ehresmann, C., Ehresmann, B. and Marquet, R. (2000) Convergence of natural and artificial evolution on an RNA loop-loop interaction: the HIV-1 dimerization initiation site. RNA, 6, 1267–1276.
16. Laugna, M. and Jette, L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downsteam of hairpin 248–271 are dispensable for dimer formation. Biochemistry, 35, 1589–1598.
17. Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. and Marquet, R. (1996) A loop-loop “kissing” complex is the essential part of the dimer linkage of genomic HIV-1 RNA. Proc. Natl. Acad. Sci. U.S.A., 93, 5572–5577.
18. Berkhour, B. and van Wamel, J.L. (1996) Role of the DIS hairpin in replication of human immunodeficiency virus type 1. J. Virol., 70, 6723–6732.
24. Mihailescu, M.R. and Marino, J.P. (2004) A proton-coupled dynamic
38. Ennifar, E., Paillart, J.C., Bernacchi, S., Walter, P., Pale, P., Decout, J.L.,
25. Muriaux, D., Fosse, P. and Paoletti, J. (1996) A kissing complex
22. Takahashi, K.I., Baba, S., Chattopadhyay, P., Koyanagi, Y.,
27. Laughrea, M. and Jette, L. (1996) HIV-1 genome dimerization:
29. Takahashi, K.I., Baba, S., Koyanagi, Y., Yamamoto, N., Takaku, H. and
35. Ulyanov, N.B., Mujeeb, A., Du, Z., Tonelli, M., Parslow, T.G. and
33. Baba, S., Takahashi, K., Noguchi, S., Takaku, H., Koyanagi, Y.,
Kawai, G. (2001) Two basic regions of NCp7 are sufficient for
50. Jossinet, F., Paillart, J.C., Westhof, E., Hermann, T., Skripkin, E.,
46. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for
57. Hagan, N.A. and Fabris, D. (2007) Dissecting the protein-RNA and
59. Marino, J.P., Gregorian, R.S. Jr, Csankovszki, G. and Crothers, D.M.
31. Aubertin, A.M., Pale, P., Dumas, P. and Marquet, R. (2006) Targeting
41. Wood, S. and Rueda, D. (2013) In: Roberts GCK (ed.) Encyclopedia of
83. 8348–8354.
Biochemistry
41. 99, 1925–1931.
101, 1189–1194.
35, 5075–5082.
41, 14762–14770.
105, 9210–9215.
7128–7139.
J.Biol.Chem., 276, 3127–31278.
J. Biol. Chem., 276, 1004–1068.
J. Biol. Chem., 31, 713–782.
Structure, 7, 1439–1449.
J. Biol. Chem., 271, 33686–33692.
Nucleic Acids Res., 41, 1502–1508.
J. Mol. Biol., 270, 5126–5133.
5126–5133.
Structure, 7, 1448–1454.
Nucleic Acids Res., 36, 7146–7156.
J. Biol. Chem., 280, 40112–40121.
J. Biol. Chem., 281, 16168–16177.
J. Biol. Chem., 280, 9445–9456.
Nucleic Acids Res., 35, 1222–1234.
Nucleic Acids Res., 36, 343–352.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 99, 507–512.
Nucleic Acids Res., 35, 1226–1234.
Nucleic Acids Res., 41, 1097–1107.
Nucleic Acids Res., 36, 9366–9374.
Nucleic Acids Res., 105, 9210–9215.
Biophys.J., 109, 2026–2034.
Nucleic Acids Res., 36, 4275–4284.
Nucleic Acids Res., 35, 771–782.
Nucleic Acids Res., 35, 1189–1194.
Nucleic Acids Res., 35, 14762–14770.
Nucleic Acids Res., 35, 96–102.
Nucleic Acids Res., 31, 5126–5133.
Nucleic Acids Res., 36, 5126–5133.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 35, 1226–1234.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 34, 2328–2339.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 7128–7139.
Nucleic Acids Res., 35, 7128–7139.
60. Lee, A.J. and Crothers, D.M. (1998) The solution structure of an RNA loop-loop complex: the CoIE1 inverted loop sequence. *Structure*, 6, 993–1005.

61. Sun, X., Zhang, Q. and Al-Hashimi, H.M. (2007) Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg2+ binding: role in the kissing-duplex structural transition. *Nucleic Acids Res.*, 35, 1698–1713.

62. Dethoff, E.A., Petzold, K., Chugh, J., Casiano-Negroni, A. and Al-Hashimi, H.M. (2012) Visualizing transient low-populated structures of RNA. *Nature*, 491, 724–728.

63. Song, R., Kafaie, J., Yang, L. and Laughrea, M. (2007) HIV-1 viral RNA is selected in the form of monomers that dimerize in a three-step protease-dependent process; the DIS of stem-loop 1 initiates viral RNA dimerization. *J. Mol. Biol.*, 371, 1084–1098.

64. Reblova, K., Fadrna, E., Sarzynska, J., Kulinski, T., Kulhanek, P., Ennifar, E., Koca, J. and Sponer, J. (2007) Conformations of flanking bases in HIV-1 RNA DIS kissing complexes studied by molecular dynamics. *Biophys. J.*, 93, 3932–3949.

65. Li, P.T. and Tinoco, I. Jr (2009) Mechanical unfolding of two DIS RNA kissing complexes from HIV-1. *J. Mol. Biol.*, 386, 1343–1356.

66. Ennifar, E., Aslam, M.W., Strasser, P., Hoffmann, G., Dumas, P. and van Delft, F.L. (2013) Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA viral genome. *ACS Chem. Biol.*, 8, 2509–2517.