

Insights on the role of nucleic acid/protein interactions in chaperoned nucleic acid rearrangements of HIV-1 reverse transcription

Hsiao-Wei Liu*, Yining Zeng*, Christy F. Landes*, Yoen Joo Kim*, Yongjin Zhu*, Xiaojing Ma*, My-Nuong Vo[†], Karin Musier-Forsyth^{†‡}, and Paul F. Barbara^{*§}

*Center for Nano and Molecular Science and Technology, University of Texas, Austin, TX 78712; and [†]Department of Chemistry, University of Minnesota, Minneapolis, MN 55455

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 25, 2006.

Contributed by Paul F. Barbara, February 7, 2007 (sent for review January 8, 2007)

HIV-1 reverse transcription requires several nucleic acid rearrangement steps that are “chaperoned” by the nucleocapsid protein (NC), including minus-strand transfer, in which the DNA transactivation response element (TAR) is annealed to the complementary TAR RNA region of the viral genome. These various rearrangement processes occur in NC bound complexes of specific RNA and DNA structures. A major barrier to the investigation of these processes *in vitro* has been the diversity and heterogeneity of the observed nucleic acid/protein assemblies, ranging from small complexes of only one or two nucleic acid molecules all the way up to large-scale aggregates comprised of thousands of NC and nucleic acid molecules. Herein, we use a flow chamber approach involving rapid NC/nucleic acid mixing to substantially control aggregation for the NC chaperoned irreversible annealing kinetics of a model TAR DNA hairpin sequence to the complementary TAR RNA hairpin, i.e., to form an extended duplex. By combining the flow chamber approach with a broad array of fluorescence single-molecule spectroscopy (SMS) tools (FRET, molecule counting, and correlation spectroscopy), we have unraveled the complex, heterogeneous kinetics that occur during the course of annealing. The SMS results demonstrate that the TAR hairpin reactant is predominantly a single hairpin coated by multiple NCs with a dynamic secondary structure, involving equilibrium between a “Y” shaped conformation and a closed one. The data further indicate that the nucleation of annealing occurs in an encounter complex that is formed by two hairpins with one or both of the hairpins in the “Y” conformation.

HIV-1 nucleocapsid protein | nucleic acid aggregation | transactivation response element DNA/RNA annealing | wild-type (WT)

During reverse transcription, an essential step in retroviral replication, single-stranded viral genomic RNA is converted into double-stranded DNA. Reverse transcription of the HIV-1 genome RNA requires two obligatory strand transfer steps. During the first strand transfer reaction, the minus-strand strong-stop DNA [(–)SSDNA] is transferred to the homologous sequence in the 3′-untranslated region of the RNA genome, by annealing (i.e., hybridization) of complementary sequences located at the 3′ end of the (–)SSDNA and genomic template, respectively (1–3). In HIV-1, the major components of (–)SSDNA transfer are the terminally redundant structured TAR elements and the nucleocapsid protein (NC) NCp7, which chaperones this annealing process and other analogous processes in reverse transcription (4–8). Fig. 1 *Upper* schematically illustrates this annealing process (9, 10). The “product” DNA/RNA duplex structure, which contains several additional Watson–Crick base pairs, is thermodynamically favored over the more weakly base-paired reactant hairpins. The D and A designation refer to optional (Donor and Acceptor) dye functionalized nucleotides used for FRET assays, and B indicates biotin-functionalized nucleotides used for immobilization purposes.

The chaperone (catalytic) activity of NC is believed to be derived from two main consequences of the nucleic-acid/protein (NC) interactions in this system. First, NC lowers the barrier for annealing by partially melting the Watson–Crick pairing of the hairpins (7, 10–14). Second, NC lowers the energy cost of bringing the hairpins together in the encounter complex by screening the negative charges of the hairpins and perhaps through specific interactions (8, 15–18). It has been extremely challenging to develop a clear understanding of the mechanistic role of the NC/nucleic acid interactions due to the tremendously diverse sets of nucleic-acid/NC complexes that have been observed for these systems *in vitro*. Single molecule spectroscopy (SMS) is uniquely capable of directly unraveling molecular structure, dynamics, and kinetics in highly heterogeneous and complex biological systems (19–22). Previously, we have shown by single-molecule FRET (SM-FRET) measurements on single immobilized transactivation response element (TAR) DNA hairpins (in the absence of TAR RNA or complementary TAR, cTAR) that NC induces a shift of the secondary structure of TAR from a fully closed conformation to a partially open, Y-shaped (i.e., “Y”) conformation, in which the L1L2 stems (i.e., bulges) are “open” or unwound and the L3L4 stems are closed (see Fig. 1 *Upper* for the L1–L4 designation) (10). In a separate study, we used SM-FRET to study the reversible annealing of various short DNA oligonucleotides to TAR to mimic the initial annealing step, demonstrating two potential sites for annealing nucleation, i.e., the 3′/5′ termini, and the hairpin loops in the L3L4 region. Annealing at the 3′/5′ termini through melting of both L1L2 has been previously denoted by the term “zipper” mechanism (9). Intermediates associated with both mechanisms were observed in the presence of NC, and the kinetics of formation of these intermediates was also measured (9).

Herein, we use SMS kinetic measurements to investigate the mechanism of the annealing kinetics of immobilized TAR hairpins to full-length complementary nonimmobilized TAR RNA or cTAR wild-type (WT) and mutant hairpins in the presence of NC. We directly observe a diverse set of annealing

Author contributions: H.-W.L. and Y. Zeng contributed equally to this work; P.F.B. designed research; H.-W.L., Y. Zeng, C.F.L., Y.J.K., Y. Zhu, and X.M. performed research; M.-N.V. and K.M.-F. contributed new reagents/analytic tools; H.-W.L., Y. Zeng, Y.J.K., Y. Zhu, and X.M. analyzed data; and P.F.B. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: SMS, single-molecule spectroscopy; TAR, transactivation response element; cTAR, complementary TAR.

[§]Present address: Department of Chemistry, Ohio State University, Columbus, OH 43210.

[§]To whom correspondence should be addressed. E-mail: p.barbara@mail.utexas.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700166104/DC1.

© 2007 by The National Academy of Sciences of the USA

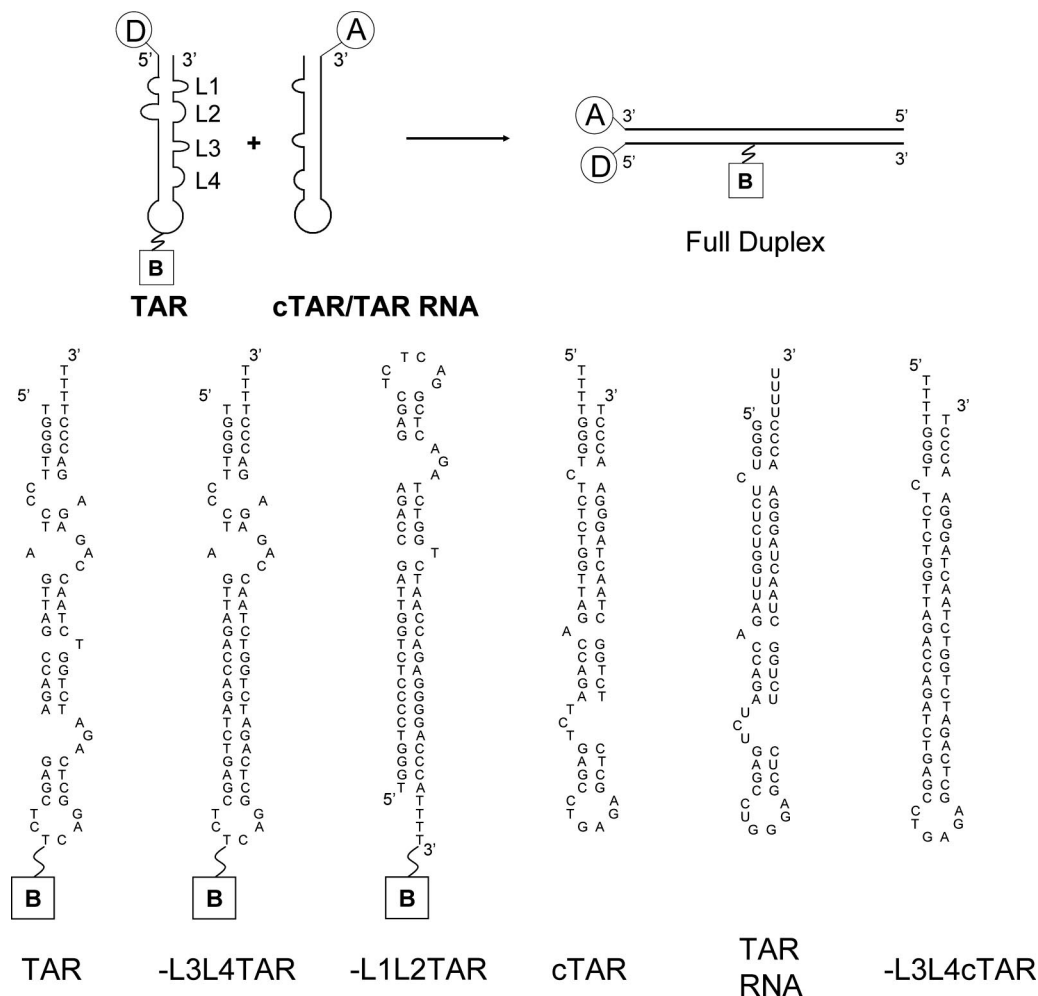


Fig. 1. Structures of various oligonucleotides used in the annealing study. The secondary structures are predicted by the mfold program (www.bioinfo.rpi.edu/applications/mfold/old/dna/).

intermediates including single NC-coated hairpins, NC-bound complexes of pairs of DNA and/or RNA hairpins, and large-scale NC/nucleic acid aggregates containing many thousand RNA, DNA, and NC molecules. This tremendous structural diversity is shown to make these reactions unsuitable for analysis by the usual “homogenous” reaction mechanistic approaches that use for example the “order” of a reaction to help identify intermediates and transition states for a reaction. By employing a combination of SMS techniques, and by simultaneously controlling large scale aggregation, we have been able to address several outstanding mechanistic issues for minus-strand transfer, including the degree of complexation and secondary structure of the reactants *in situ*, the identity of the site of the nucleation of annealing, and whether the so-called “kissing loop” interaction (23–27) is an important interaction in the mechanism.

Results

Annealing reactions were investigated between nearly all of the pair-wise combinations of immobilized and nonimmobilized hairpins that are shown in Fig. 1, including the HIV-1 WT pair, TAR/TAR RNA. The various hairpins investigated include various mutants that were designed to probe how the kinetics depends on the presence or absence of internal bulges in the hairpins. Unless otherwise noted, the annealing experiments used NC concentrations = 890 nM and nucleotide/NC molar

concentration ratios <5 , ensuring saturation binding of NC to the hairpins (28–30).

Immobilization of the TAR hairpins on a biologically compatibilized cover-slip located in a flow chamber with the biotin-streptavidin approach (Fig. 1) was used (9) to keep one of the annealing reactants stationary and thereby allow for single-molecule spectroscopic microscopy measurements at specific brief time intervals over the duration of the reaction, i.e., typically hundreds to thousands of seconds. The experimental design used a multiple syringe pump flow system allowing for rapid mixing of the NC with nonimmobilized hairpin and then rapid delivery of this solution to the microscopy cell (see *Materials and Methods*). This procedure effectively suppressed formation of large NC/nucleic acid aggregates (in most cases). Aggregation has been reported to be a serious obstacle to making direct *in vitro* measurements of the annealing kinetics for NC structures that contain the WT N-terminal 3_{10} helix domain (24, 31). In addition to DNA/RNA annealing (as in HIV-1), various DNA/DNA annealing reactions were also investigated as a mechanistic comparison and to take advantage of the easier access to DNA mutants and the smaller tendency of DNA to aggregate with NC (32–34).

The immobilized hairpins were labeled at the 5' end with a Cy3 dye (FRET donor dye) and the complementary nonimmobilized hairpins were labeled at the 3' end with a Cy5 acceptor (A), with the usual attachment configuration (9). Cy3/Cy5 is a well estab-

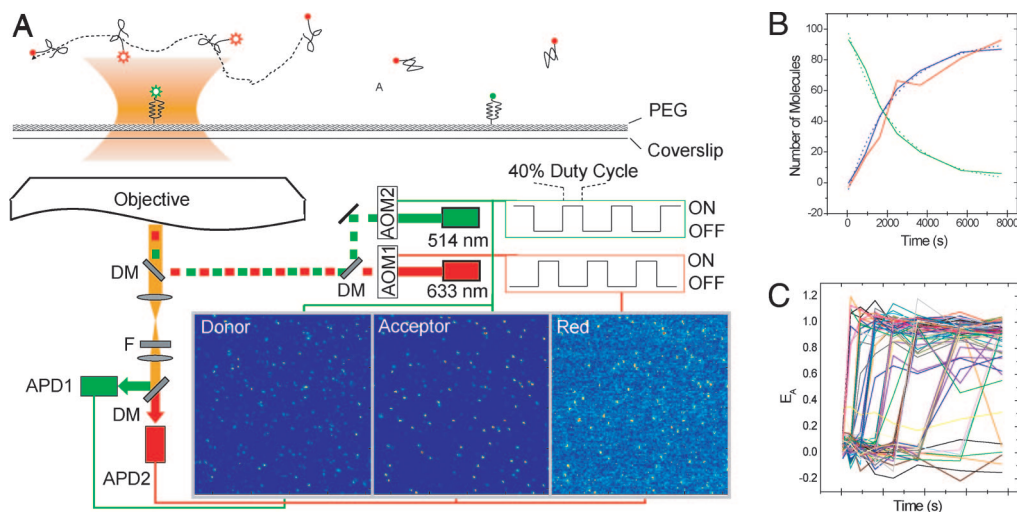


Fig. 2. A typical single molecule kinetic measurement of TAR and 2 nM cTAR annealing reaction at 0.2 mM Mg^{2+} and 889 nM NC in buffer solution. (A) The alternating two-color, laser excitation microscopy setup. (B) The number of molecules in three channels (donor, acceptor, and red) during the course of the annealing reaction. (C) E_A of each single molecule as a function of time, where each colored line corresponds to a single molecule. AOM, acoustic optical modulator; DM, dichroic mirror; F, notch filter; APD, avalanche photodiode; PEG, poly(ethylene glycol). Modulators controlled by two 180° out-of-phase square-wave signals give two-color alternating-laser excitation. After being filtered by a notch filter, the fluorescence is detected by APD1 and -2. Donor and acceptor channel fluorescence detected by APD1 and -2, respectively, is counted by two counters while the green excitation laser is on. Red channel fluorescence is also detected on APD2, except that it is counted by a third counter while red excitation laser is on.

lished donor/acceptor dye pair allowing for an instantaneous measurement of the interhairpin Cy3–Cy5 distance by SM-FRET (19, 35, 36). Irreversible annealing kinetics were initiated by exposing a dilute immobilized sample of the Cy3-TAR to a “fresh” solution of Cy5–cTAR and NC at zero time, $t = 0$, in analogy to a stopped flow experiment (see Fig. 2A Top) as shown in Fig. 2 for the NC catalyzed annealing of TAR to cTAR DNA. The fluorescence spots in the confocal images for the Donor and Acceptor channels were used to determine the acceptor and donor intensities, $I_A(t)$ and $I_D(t)$, and the apparent FRET efficiency, E_A , defined as:

$$E_A(t) = \frac{I_A(t)}{I_A(t) + I_D(t)}. \quad [1]$$

By measuring E_A for each hairpin at various times, t , after introducing the complementary Cy5–hairpin solution, FRET trajectories (Fig. 2C) were recorded to monitor the instantaneous distance between the 5′ end of the immobilized Cy3–TAR hairpin and the 3′ end of the Cy5 hairpins. The time spacing between FRET points (i.e., confocal images) was varied during the experiment to minimize photobleaching. Typically, the time spacing was >120 s, i.e., the time to record one confocal image. In certain cases, 10-ms resolved FRET trajectories were recorded for individual hairpins at specific times along the course of the reaction by parking the confocal apparatus on a specific hairpin.

In the presence of both NC and WT complementary hairpins in solution, the observed E_A trajectories exhibit discrete jumps from an E_A value near zero to a value near unity (Fig. 2C) at various times during the annealing reaction. With either or both of NC and complementary hairpins absent, the E_A value does not vary from zero over the entire time scale of the experiments (data not shown). The discrete, “telegraphic” nature of the annealing trajectories is especially apparent in “kinetic” histograms of the SMS E_A determined for specific time windows after the initiation of the reaction for an ensemble of reacting hairpins, see Fig. 3. Thus, the SMS kinetic results directly reveal, in a way that would not be possible with ensemble measurements, that the annealing reaction evolves to kinetic “stable-states,” i.e., all

other nucleic acid rearrangements must occur on a more rapid time-scale. The mean FRET value ($\langle E_A \rangle$), as a function of reaction time for an ensemble of immobilized WT and mutant hairpins is shown by square data points in Fig. 4A.

The number of reactant and product molecules, N_R and N_P , respectively, were counted directly by using a FRET threshold of 0.4 to distinguish between reactants and products (see Fig. 2B). These N_R and N_P data for TAR/cTAR annealing under saturated NC binding (green and blue solid lines, respectively) were well-fit with a single exponential function (dashed lines), which is the predicted kinetic behavior for an irreversible bimolecular reaction under the present pseudo-first-order conditions for these nonimmobilized hairpins. The best-fit parameters include a pseudo-first-order rate constant, k_{app} , and a percent annealed, the latter of which is close to the expected 100%. The single exponential behavior suggests the existence of a kinetic bottle-

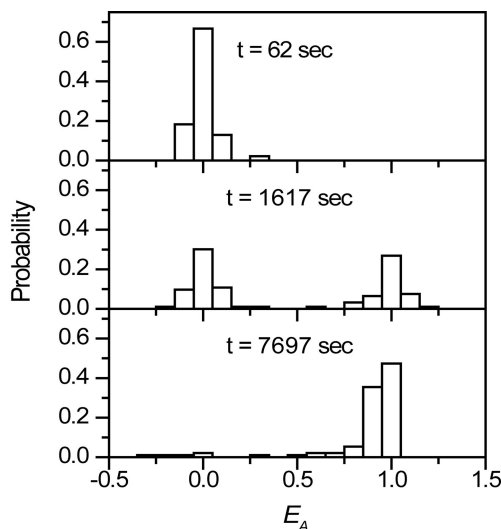


Fig. 3. $E_A(t)$ histograms constructed at different time after initiation of the NC-induced annealing of Cy3-TAR annealing with Cy5-cTAR.

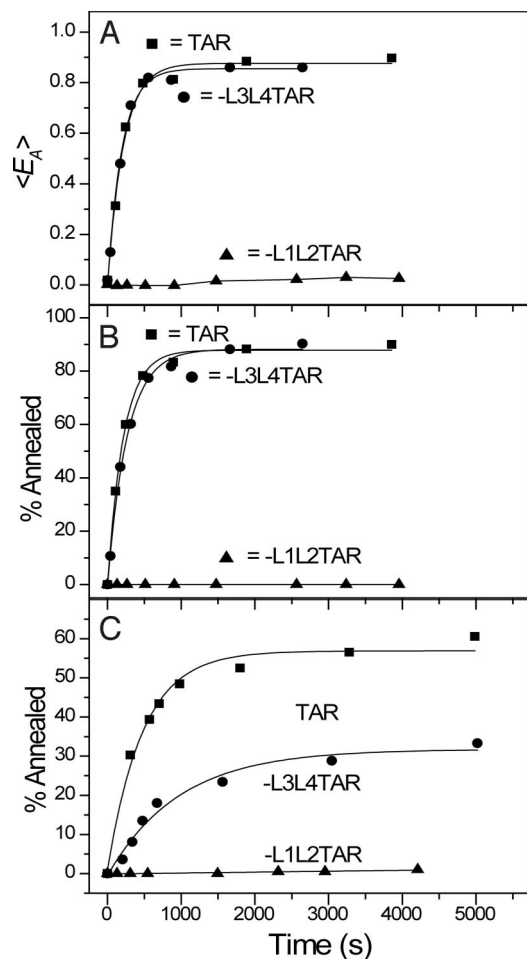


Fig. 4. SM-FRET measurements of the annealing kinetics. $\langle E_A \rangle$ (A) and %annealed (B) vs. time for the annealing of cTAR with WT and mutant TAR DNA hairpins at 10 nM cTAR, 2 mM Mg^{2+} and 890 nM NC. (C) %Annealed vs. time for annealing of TAR RNA with various immobilized WT and mutant TAR DNA hairpins at 0.2 mM Mg^{2+} and 890 nM NC. For TAR DNA, 10 nM TAR RNA was used, whereas for -L3L4 TAR and -L1L2TAR, 25 nM TAR RNA was used.

neck for annealing. In chemical kinetics, a single bottleneck often implies a single well defined transition state. In the case of the annealing reaction, the conclusion of a well defined bottleneck applies to the aggregation free (homogenous solution) mode of the annealing reaction. A more complex mechanism that involves more than one kinetic bottleneck applies when aggregation is not controlled (see below).

A distinct advantage of the scanning confocal method for SMS is that it allows for synchronous detection of the Cy3 and Cy5 fluorescence intensity while rapidly switching the laser excitation wavelength between 514 nm and 633 nm, which are the wavelengths that excite Cy3 and Cy5, respectively. As shown in Fig. 2, the 633-nm excitation induces only Cy5 fluorescence (denoted by red), which is a quantitative measure of the number of bound Cy5-labeled hairpins at various times during the reaction, after calibration. The SMS kinetic data in Fig. 2B shows that the number of associated cTAR hairpins of any type (red curve) is equal to the number of annealed hairpins (blue curve) throughout the annealing reaction, within experimental error. This finding demonstrates that cTAR is not significantly associated with the TAR until annealing has occurred. A quantitative analysis of data for a subensemble of ≈ 100 unannealed immobilized TAR hairpins in the presence of cTAR leads to a value of >70 nM for the apparent dissociation constant of the un-

annealed TAR/cTAR complex [see supporting information (SI) Text for detailed information on this analysis]. The same type of experiments with the other immobilized hairpin mutants, shown in Fig. 1, gave analogous results. An exception to this behavior occurs during large scale NC induced nucleic acid aggregation (see below).

Further information on the secondary structure and degree of aggregation of the reactant TAR hairpins was obtained by recording the E_A histogram of immobilized doubly labeled 5'/3' donor/acceptor TAR in the presence of both NC and the unlabeled complementary cTAR reactant. During the reaction the E_A histogram exhibited a major peak at 0.8 assigned to the partially open "Y" conformation (9, 10), and a second peak at $E_A \approx 0$ assigned to the duplex product of annealing, which has a significantly increased spacing between the 3' and 5' ends (data not shown). The presence of nonlabeled cTAR under typical reaction (10 nM cTAR) conditions has no measurable effect on the E_A 0.8 peak assigned to "Y" (data not shown), further demonstrating that the reactant TAR DNA hairpins are not significantly associated with cTAR under these conditions. Thus, a single "Y" hairpin coated by multiple NCs is the dominant form for the TAR DNA reactant in the NC-induced annealing mechanism of TAR DNA.

NC/Nucleic Acid Aggregation. Under conditions in which large-scale NC/nucleic acid aggregates are present, e.g., at high cTAR (or TAR RNA) concentration, the observed rate constants for NC-induced annealing for the various hairpins fluctuated greatly from trial to trial. This fluctuation is reflected in Table 1 by the extraordinarily large standard deviations for TAR/cTAR annealing at high cTAR concentrations. A number of hypothetical origins for the rate fluctuations, such as sticky syringes, temperature variations, and surface adsorption of NC or the nucleic acids were ruled out by careful controls. Ultimately, the reaction rate fluctuations were assigned to the aggregation of the cTAR with NC in large molecular aggregates that contain hundreds to thousands of cTAR and NC molecules. The number of cTAR molecules per aggregate was estimated from the fluorescence burst intensity for individual aggregates, using the known fluorescence intensity for individual hairpins at the same excitation intensity as a calibration.

Direct evidence for NC-induced aggregation was obtained by recording time resolved fluorescence intensity curves, $I(t)$, for Cy5-labeled cTAR in solution during an annealing trial with the labeling pattern indicated by the top equation in Fig. 1. Here, the scanning stage was positioned over a region that was free of immobilized hairpins, and the fluorescence induced by 633-nm excitation was recorded for a relatively high concentration of Cy5-cTAR in solution in the excited volume of the confocal, in the absence (blue) and presence (red) of NC (see Fig. 5). In the absence of NC, $I(t)$ only exhibits short time fluctuation due to photon shot-noise and concentration fluctuations that are consistent with the diffusion of single, nonaggregated Cy5-cTAR hairpins. In contrast, in the presence of NC, the $I(t)$ curves show clear evidence of large scale nucleic acid aggregation. For example, intense blips are observed due to single aggregates containing thousands of hairpins. Also, a significant decrease in the steady-state concentration of the cTAR monomer is observed, resulting from incorporation of cTAR in the NC-assembled aggregates. The $I(t)$ curves indicate that both hairpin monomers and aggregated hairpins are present simultaneously. Aggregation was observed to be especially severe for solutions with high hairpin concentrations, and was more severe for RNA than DNA. These various observations suggest that the aggregation process is analogous to a nucleation-controlled precipitation.

By using high flow rates (freshly mixed solutions), it was possible to suppress the formation of aggregates in the presence of NC for small nucleic acid concentrations. This suppression

Table 1. The apparent second-order TAR/cTAR annealing rate constants, k_a , at various NC, Mg^{2+} , and cTAR concentrations

[Mg^{2+}], mM	[NC], nM	[cTAR], nM	k_a , $M^{-1} s^{-1}$	[NC]/[nt]
2	890	5	$(3.64 \pm 0.34) \times 10^5$	2.8
		10	$(2.72 \pm 1.2) \times 10^5$	1.4
		20	$(3.21 \pm 0.31) \times 10^5$	0.7
1	890	10	$(6.66 \pm 0.47) \times 10^5$	1.4
0.2			$(8.37 \pm 0.048) \times 10^5$	1.4
0.2	500	50	$(1.69 \pm 0.09) \times 10^5$	0.16
		83.3	$(9.16 \pm 10.9) \times 10^2$	0.09
0.2	300	50	$(3.73 \pm 2.77) \times 10^4$	0.09
	<200		No annealing	<0.063

The kinetics were measured under pseudo first order conditions, where the $[cTAR] \gg [TAR]$. The apparent second-order rate constant, k_a was estimated by dividing the pseudo-first-order rate constant by $[cTAR]$. In fact, the reaction is only second-order at high values of $[NC]/[nt]$ (see text for further detail). The reaction conditions are: buffer A (40 mM NaCl, 25 mM Hepes, pH 7.3 and glucose oxygen scavenger system) at room temperature. k_a is the average of at least three trials for each reaction.

was confirmed by fluorescence correlation spectroscopy (FCS) on Cy5-cTAR in the absence and presence of NC (Fig. 5 *Inset*). The FCS data in both cases are well fit by the standard model with best-fit diffusion constants that are consistent with expectations for a single hairpin (see *SI Text*). The aggregation phenomenon was observed to produce huge fluctuations of the annealing rate at high cTAR and TAR RNA concentrations, even inducing a “stalling effect,” i.e., no detectable reaction despite high cTAR and NC concentration (see Table 1). Over this range, the kinetic order of the reaction for cTAR was observed to vary from first-order at <20 nM cTAR concentrations, to zero-order at concentrations in the 25–50 nM range, and then ultimately to actually exhibit negative orders when the reaction stalled at high cTAR concentration.

Annealing Rate Trends in the Absence of Aggregation. Aggregation, which appears to be a kinetically controlled process, can be largely avoided for low cTAR and TAR RNA concentrations (i.e., <10

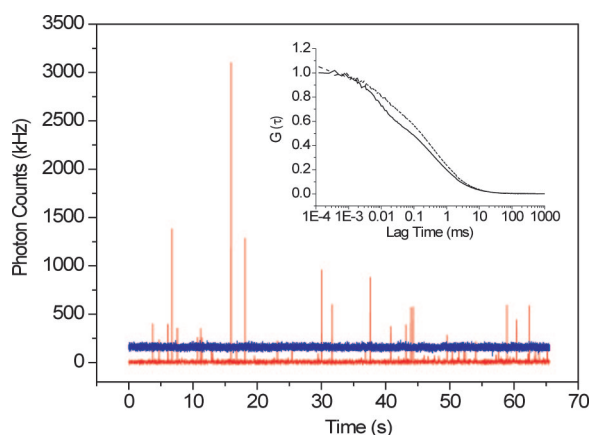


Fig. 5. Representative $I(t)$ curves detecting the emission of Cy5-cTAR DNA. Blue lines are the trajectory of Cy5-cTAR DNA (50 nM) without NC present, exhibiting a continuous signal of ≈ 160 kHz (detected photons per second). Red lines correspond to Cy5-cTAR DNA (50 nM) with NC (500 nM), revealing intense blips due to aggregation and a decrease in the steady-state signal from monomer Cy5-cTAR DNA to ≈ 4 kHz. (*Inset*) FCS curves of cTAR (5 nM) (solid line) and cTAR (5 nM) with NC (500 nM) (dashed line) showing increase of τ_D due to NC binding to cTAR. Here, τ_D is the diffusion time for cTAR or cTAR/NC complexes passing through the focal volume of the laser.

nM). Under these conditions, analysis of WT and mutant TAR hairpins reveals several mechanistically informative trends (see the data in Tables 1 and 2 and Fig. 4). The WT TAR/cTAR and TAR/TAR RNA (with L1–L4 preserved for both TAR and cTAR/TAR RNA) reactions are observed to be overall second-order, i.e., the pseudo-first-order rate vs. cTAR is linearly proportional to $[cTAR]$ (see Table 1). At low NC concentrations, the rate falls off precipitously with $[NC]$, as expected for NC-induced cooperative melting of TAR and cTAR, which is a key step in the annealing process. For -L1L2TAR/TAR annealing, the data clearly show that the L1L2 bulges of TAR are required (and in fact are sufficient for rapid annealing, see Fig. 4*A* and *B*). In contrast, removing the L3L4 bulges has little kinetic consequence. Similar trends are observed for TAR/TAR RNA annealing. Although removing the L3L4 bulges does not stop the rapid initiation of annealing, it apparently shifts the annealing equilibrium toward reactants. The lower apparent annealing percentage for TAR RNA has been extensively investigated and has been assigned to several factors, including incomplete dye labeling (probably $\approx 70\%$), errors in the synthesis producing a small concentration of TAR RNA hairpins with missing bases, and complications due to aggregation.

Discussion

Based on the information obtained herein, a refined and more specific mechanism for the NC chaperoned annealing of TAR to cTAR or TAR RNA has been constructed, as shown in Fig. 6. This mechanism applies exclusively to the annealing reaction in the absence of large scale nucleic acid/protein aggregates. The SMS results demonstrate that the TAR reactant is predominantly a single NC-coated hairpin with a dynamic secondary structure, involving a partially open “Y” shaped conformation, as shown in Fig. 6*i*. The FCS data demonstrate that the complementary cTAR or TAR RNA reactants are single nonaggregated hairpins, i.e., except for the bound NCs. It is reasonable, based on the experiments for TAR, to assume that the cTAR or TAR RNA reactants also have a partially melted secondary structure, although future experiments will be necessary to verify this conjecture.

The mechanism portrayed in Fig. 6 assumes that rapid and reversible association and partial melting precedes a much slower nucleation event. Our observation that the L1L2 bulges of TAR are required (and in fact are sufficient) for rapid annealing, coupled with the previous observation that hairpins missing the L1L2 bulge region do not undergo NC-induced melting, strongly suggests that the encounter complex is formed by two hairpins with one or both of the hairpins in the “Y” conformation (10, 37). This conclusion is reflected in the first step in Fig. 6*iii*). The rate-limiting-step (RLS) for annealing, furthermore, is hypothesized to be local annealing at specific locations along the hairpins, which can “nucleate” the annealing process. A steady-state solution of this mechanism predicts the second-order overall annealing kinetics (in nonaggregating conditions), which is in agreement with the experimental results. Furthermore, the close similarity of the annealing rate constant of TAR with cTAR, to that for other oligonucleotides that should strongly favor the zipper mechanism (e.g., -L3L4TAR and a short oligonucleotide targeted for L1L2 region; ref. 9) strongly suggests that the zipper mechanism dominates the annealing reaction.

We have previously shown that a short, locally targeted oligonucleotide for the hairpin loop, including the L3L4 regions, is capable of annealing, although at a slower rate than the zipper nucleation route (9). This finding implies that there may be two pathways for annealing for full length cTAR and TAR RNA. However, for the full-length oligonucleotides, nucleation by either route ultimately leads to the same duplex annealed product. The similar rates of the loop and zipper routes (only an order of magnitude slower at low Mg^{2+} concentrations) suggests that a

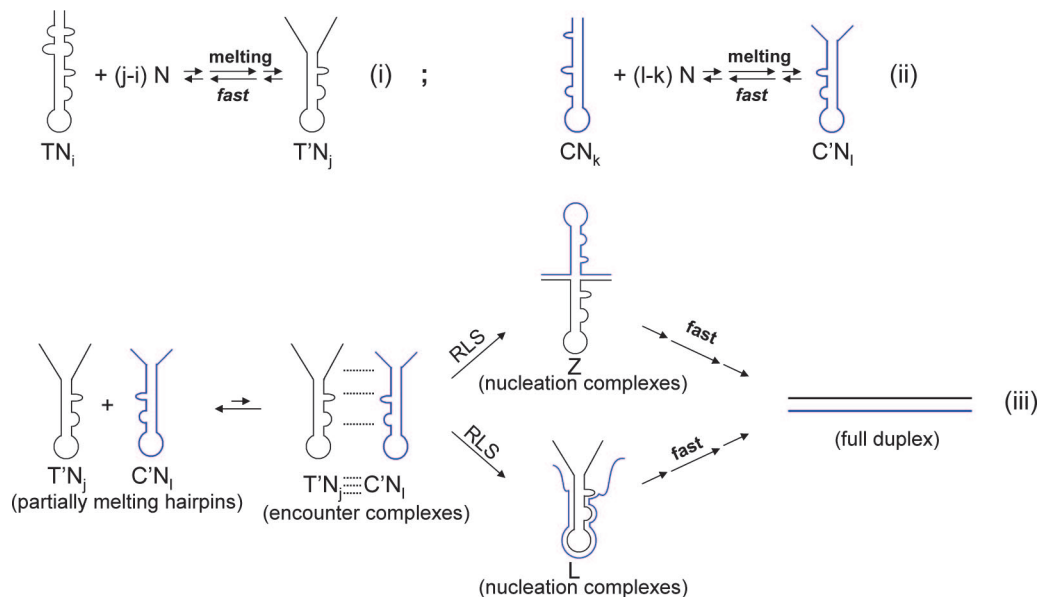


Fig. 6. The hypothetical kinetic scheme of TAR DNA annealing with its complements chaperoned by HIV-1 NC. Here, T denotes TAR DNA, C denotes complementary cTAR DNA or TAR RNA, and N denotes NC. In this scheme, N bound to T and C leads to a partially melting structure, namely Y form of T (T') and C (C'). The subscripts, i , j , k , and l are used to describe the number of NC bound to nucleotides. Two partially melting hairpins form an encounter complex that leads to the formation of nucleation complexes. The annealing can go through either zipper nucleation or loop nucleation, therefore, forming zipper nucleation complexes (Z) or loop nucleation complexes (L) that leads to the formation of fully annealed duplexes.

common mechanism may be operating in all cases. Nucleation of annealing at two different locations corresponds to a “bifurcation” of the reaction path and technically corresponds to two distinct transition states (and mechanisms) for the reaction. It should be emphasized, however, that the two pathways are highly analogous, involving nucleation of annealing in locally single-strand regions in the NC melted and associated complex of the two hairpins. Thus, a better description for the reaction than two-transition states may be a reaction mechanism in which the transition state is a broad multidimensional region along a “wide” reaction path with two or more adjacent forms of the same type of transition states corresponding to different nucleation sites. Such complexities are not surprising in nucleic acid arrangements due to the rugged energy landscape (6, 38, 39).

The possibility that annealing can be initiated by nucleation through the hairpin loops (L3L4 region) has been recognized for some time, and has been previously assumed to occur through a kissing hairpin loop interaction mechanism, which is distinct from that summarized in Fig. 6 (24, 40, 41). The key feature of the kissing mechanism is an encounter complex that is stabilized by base pairing between the terminal loop region, i.e., the so-called kissing-loop interaction. Although the present set of experiments do not directly probe the possible involvement of kissing-loop encounter complexes, the observations that TAR/L3L4 nucleated annealing is as rapid as WT annealing indicates that kissing loop-initiated annealing is not necessary for rapid annealing.

The SMS results strongly suggest that the common approach of investigating the NC-induced annealing kinetics of nucleic acids with aggregates present must be undertaken with great caution due to the exceedingly complex and heterogeneous nature of these reactions at the molecular level. Nevertheless, the annealing reaction kinetics and mechanism within large scale aggregates is of interest due to its potential similarity with certain aspects of the annealing process *in vivo*. Indeed, future studies on annealing in large scale aggregates may shed light on how the spatial organization within complex NC/nucleic acid assemblies

controls the rates and pathways of annealing in the reverse transcription mechanism of HIV-1.

Conclusions

Herein, by using a flow system with rapid mixing, we have been able to use a FRET-based approach to investigate the aggregation free annealing kinetics with WT NC. Previously, to suppress aggregation effects on the annealing kinetics it was necessary to use a truncated NC (12–55) peptide that lacks the so-called aggregating terminal domain of NC (31). Insights into the *in vitro* annealing mechanisms in the minus-strand transfer step of HIV-1 reverse transcription have been obtained. Specifically, the NC chaperoned irreversible annealing kinetics of a model TAR DNA hairpin sequence to the complementary cTAR (or TAR RNA) hairpin to form an extended duplex was investigated by SMS kinetic approaches. Specific information on the composition and secondary

Table 2. The effect of loop removal on the second-order rate constants, k for TAR DNA:cTAR annealing in the presence of 890 nM NC and 2 mM Mg²⁺

Nonimmobilized DNA	Immobilized DNA	k_a (M ⁻¹ s ⁻¹), average \pm SD*
cTAR	TAR [†]	$(2.72 \pm 1.2) \times 10^5$
	-L3L4TAR	$(3.03 \pm 0.61) \times 10^5$
	-L1L2TAR	No annealing
-L3L4cTAR	TAR	$(2.16 \pm 0.96) \times 10^5$
	-L3L4TAR	$(1.98 \pm 0.71) \times 10^5$

The reactions were run in buffer A (40 mM NaCl, 25 mM Hepes, pH 7.3, and glucose oxygen scavenger system) at room temperature. The apparent second-order rate constant, k_a was estimated by dividing the pseudo-first-order rate constant by the [cTAR]. k_a is the average of at least three trials for each reaction.

*For -L3L4TAR/cTAR annealing, a biexponential fitting curve can be obtained, which gives a fast rate, $(3.52 \pm 0.75) \times 10^5$ (M⁻¹s⁻¹), and a slow rate, $(5.33 \pm 2.71) \times 10^4$ (M⁻¹s⁻¹). A single exponential fitting is shown.

[†]For TAR/cTAR annealing at different concentrations of cTAR and Mg²⁺ (see Table 1 for the details).

structures of key intermediates and transition states was obtained by using SMS fluorescence tools (FRET, molecule counting, and correlation spectroscopy) in combination with a flow chamber approach involving rapid NC/nucleic acid mixing to substantially control aggregation. The SMS results demonstrate that the TAR hairpin reactant is predominantly a single NC-coated hairpin with a dynamic secondary structure, involving equilibrium between a “Y” shaped conformation and a closed one. The data further suggest that the nucleation of annealing occurs in an encounter complex that is formed by two hairpins with one or both of the hairpins in the “Y” conformation.

Materials and Methods

Experimental Methods. SMS data in this paper were recorded by a home-built sample-scanning confocal microscope with separate detection channels for detecting Cy3 and Cy5 emission (10). The Cy3 and Cy5 fluorescence intensity was synchronously detected while rapidly switching the laser excitation between 514 nm and 633 nm, which selectively excited Cy3 and Cy5, respectively. Spatial analysis of the confocal images allowed for colocalization of the Cy3- and Cy5-labeled hairpins. This procedure also yielded the intensity of the “free” Cy5-TAR in the flowing solution above the immobilized hairpins (which is detected as a 633 nm excited constant background in the images). By monitoring the Cy5 intensity resulting from direct excitation at 633

nm, it was possible to measure the number of Cy5-labeled hairpins, N_{cTAR} , associated (but not necessarily annealed) with a specific immobilized Cy3-TAR at various times during the annealing reaction.

Sample Preparation. HIV-1 NC was synthesized as described (9). Functionalized DNA hairpins (purchased from TriLink Biotechnologies, San Diego, CA) and RNA hairpins (purchased from Dharmacon RNA Technologies, Lafayette, CO) were used without further purification as described (9). Three syringe pumps delivered three solutions, NC, nonimmobilized hairpins (containing Mg^{2+}), and buffer A (containing Mg^{2+} as well), separately. The rapid mixing of solutions was achieved by injecting solutions at 10 μ l/min for at least 10 min, and then reducing the flow rate to \approx 1 μ l/min for the remaining reaction. Acquisition of images and the solution injection were started at the same time. All of the solutions contained buffer A (40 mM NaCl, 25 mM Hepes, pH 7.3 and glucose oxygen scavenger system; ref. 9).

We thank Drs. George Barany, Daniel G. Mullen, and Ms. Brandie Kovaleski (all from University of Minnesota, Minneapolis) for chemical synthesis of NC. This work was supported by National Institutes of Health (NIH) Grants GM65818 (to P.F.B.) and GM65056 (to K.M.-F.), NIH postdoctoral National Research Service Award GM073534 (to C.F.L.), and the Welch Foundation (P.F.B.).

- Luo GX, Taylor J (1990) *J Virol* 64:4321–4328.
- Gilboa E, Mitra SW, Goff S, Baltimore D (1979) *Cell* 18:93–100.
- Peliska JA, Benkovic SJ (1992) *Science* 258:1112–1118.
- Tsuchihashi Z, Brown PO (1994) *J Virol* 68:5863–5870.
- Darlix J-L, Lapadat-Tapolsky M, de Rocquigny H, Roques BP (1995) *J Mol Biol* 254:523–537.
- Herschlag D (1995) *J Biol Chem* 270:20871–20874.
- Rein A, Henderson LE, Levin JG (1998) *Trends Biochem Sci* 23:297–301.
- Levin JG, Guo J, Rouzina I, Musier-Forsyth K (2005) *Prog Nucleic Acid Res Mol Biol* 80:217–286.
- Liu H-W, Cosa G, Landes CF, Zeng Y, Kovaleski BJ, Mullen DG, Barany G, Musier-Forsyth K, Barbara PF (2005) *Biophys J* 89:3470–3479.
- Cosa G, Harbron EJ, Zeng Y, Liu H-W, O'Connor DB, Eta-Hosokawa C, Musier-Forsyth K, Barbara PF (2004) *Biophys J* 87:2759–2767.
- Hong MK, Harbron EJ, O'Connor DB, Guo J, Barbara PF, Levin JG, Musier-Forsyth K (2003) *J Mol Biol* 325:1–10.
- Bernacchi S, Stoylov S, Piemont E, Ficheux D, Roques BP, Darlix J-L, Mely Y (2002) *J Mol Biol* 317:385–399.
- Beltz H, Azoulay J, Bernacchi S, Clamme J-P, Ficheux D, Roques B, Darlix J-L, Mely Y (2003) *J Mol Biol* 328:95–108.
- Azoulay J, Clamme J-P, Darlix J-L, Roques BP, Mely Y (2003) *J Mol Biol* 326:691–700.
- Le Cam E, Coulaud D, Delain E, Petitjean P, Roques BP, Gerard D, Stoylova E, Vuilleumier C, Stoylov SP, Mely Y (1998) *Biopolymers* 45:217–229.
- Stoylov SP, Vuilleumier C, Stoylova E, De Rocquigny H, Roques BP, Gerard D, Mely Y (1997) *Biopolymers* 41:301–312.
- Dibhajji F, Khan R, Giedroc DP (1993) *Protein Sci* 2:231–243.
- Lapadat-Tapolsky M, De Rocquigny H, Van Gent D, Roques B, Plasterk R, Darlix JL (1993) *Nucleic Acids Res* 21:831–839.
- Ha T, Rasnik I, Cheng W, Babcock HP, Gauss G, Lohman TM, Chu S (2002) *Nature* 419:638–641.
- Zhuang X, Bartley LE, Babcock HP, Russell R, Ha T, Herschlag D, Chu S (2000) *Science* 288:2048–2051.
- Kim HD, Nienhaus GU, Ha T, Orr JW, Williamson JR, Chu S (2002) *Proc Natl Acad Sci USA* 99:4284–4289.
- Xie Z, Srividya N, Sosnick TR, Pan T, Scherer NF (2004) *Proc Natl Acad Sci USA* 101:534–539.
- Chang KY, Tinoco I (1997) *J Mol Biol* 269:52–66.
- Vo M-N, Barany G, Rouzina I, Musier-Forsyth K (2006) *J Mol Biol* 363:244–261.
- Haddrick M, Lear AL, Cann AJ, Heaphy S (1996) *J Mol Biol* 259:58–68.
- Paillart J-C, Skripkin E, Ehresmann B, Ehresmann C, Marquet R (1996) *Proc Natl Acad Sci USA* 93:5572–5577.
- Paillart J-C, Westhof E, Ehresmann C, Ehresmann B, Marquet R (1997) *J Mol Biol* 270:36–49.
- Fisher RJ, Rein A, Fivash M, Urbaneja MA, Casas-Finet JR, Medaglia M, Henderson LE (1998) *J Virol* 72:1902–1909.
- Karpel RL, Henderson LE, Oroszlan S (1987) *J Biol Chem* 262:4961–4967.
- Urbaneja MA, Kane BP, Johnson DG, Gorelick RJ, Henderson LE, Casas-Finet JR (1999) *J Mol Biol* 287:59–75.
- Godet J, de Rocquigny H, Raja C, Glasser N, Ficheux D, Darlix J-L, Mely Y (2006) *J Mol Biol* 356:1180–1192.
- Vuilleumier C, Bombarda E, Morellet N, Gérard D, Roques BP, Mély Y (1999) *Biochemistry* 38:16816–16825.
- Paoletti AC, Shubsda MF, Hudson BS, Borer PN (2002) *Biochemistry* 41:15423–15428.
- Maki AH, Ozarowski A, Misra A, Urbaneja MA, Casas-Finet JR (2001) *Biochemistry* 40:1403–1412.
- McKinney SA, Declais A-C, Lilley DMJ, Ha T (2003) *Nat Struct Biol* 10:93–97.
- Rasnik I, Myong S, Cheng W, Lohman TM, Ha T (2004) *J Mol Biol* 336:395–408.
- Cosa G, Zeng Y, Liu H-W, Landes CF, Makarov DE, Musier-Forsyth K, Barbara PF (2006) *J Phys Chem B* 110:2419–2426.
- Russell R, Herschlag D (2001) *J Mol Biol* 308:839–851.
- Russell R, Zhuang X, Babcock HP, Millett IS, Doniach S, Chu S, Herschlag D (2002) *Proc Natl Acad Sci USA* 99:155–160.
- Berkhout B, Vastenhouw NL, Klasens BIF, Huthoff H (2001) *RNA* 7:1097–1114.
- Kanevsky I, Chaminade F, Ficheux D, Moumen A, Gorelick R, Negroni M, Darlix J-L, Fosse P (2005) *J Mol Biol* 348:1059–1077.
- Krichevsky O, Bonnet G (2002) *Rep Prog Phys* 65:251–297.