BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Comparative Structural Effects of HIV-1 Gag and Nucleocapsid Proteins in Binding to and Unwinding of the Viral RNA Packaging Signal

Neil M. Bell,[†,‡] Julia C. Kenyon,[†] Shankar Balasubramanian,[‡,§] and Andrew M. L. Lever*[,†]

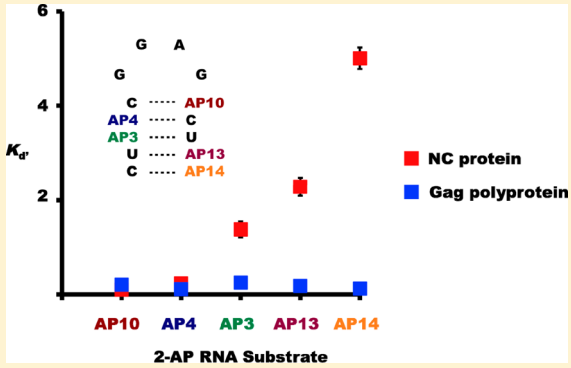
[†]Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 OQQ, U.K.

[‡]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

[§]Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, U.K.

Supporting Information

ABSTRACT: The major RNA binding region of the HIV-1 Gag polyprotein is the nucleocapsid (NC) domain, which is responsible for the specific capture of the genomic RNA genome during viral assembly. The Gag polyprotein has other RNA chaperone functions, which are mirrored by the isolated NC protein after physiological cleavage from Gag. Gag, however, is suggested to have superior nucleic acid chaperone activity. Here we investigate the interaction of Gag and NC with the core RNA structure of the HIV-1 packaging signal (Ψ), using 2-aminopurine substitution to create a series of modified RNAs based on the Ψ helix loop structure. The effects of 2-aminopurine substitution on the physical and structural properties of the viral Ψ were characterized. The fluorescence properties of the 2-aminopurine substitutions showed features consistent with the native GNAR tetraloop. Dissociation constants (K_d) of the two viral proteins, measured by fluorescence polarization (FP), were similar, and both NC and Gag affected the 2-aminopurine fluorescence of bases close to the loop binding region in a similar fashion. However, the influence of Gag on the fluorescence of the 2-aminopurine nucleotides at the base of the helix implied a much more potent helix destabilizing action on the RNA stem loop (SL) versus that seen with NC. This was further supported when the viral Ψ SL was tagged with a 5′ fluorophore and 3′ quencher. In the absence of any viral protein, minimal fluorescence was detected; addition of NC yielded a slight increase in fluorescence, while addition of the Gag protein yielded a large change in fluorescence, further suggesting that, compared to NC, the Gag protein has a greater propensity to affect RNA structure and that Ψ helix unwinding may be an intrinsic step in RNA encapsidation.



The major structural protein of retroviruses, including HIV-1, is the Gag polyprotein, which can form virus-like particles when expressed alone in cells.[1] Gag has four major domains into which it is cleaved during viral maturation, a process associated with viral budding from the cell.[2,3] In HIV, one of these four domains has two zinc knuckle motifs and on cleavage forms the nucleocapsid protein (NC) that coats the RNA genome within the mature virion.[2,4] The RNA chaperone functions of NC have been extensively studied;[5−9] however, the ability of NC to fully unwind nucleic acid helices is not certain. Early circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopic evidence for binding of NC to tRNA suggested that NC alters base stacking interactions, but it was unclear whether unwinding of the nucleic acid helices was occurring.[10−12] Musier-Forsyth et al. used several techniques to monitor the binding of NC to tRNA, and although helix unwinding of the tRNA was not observed, there was significant unstacking of the bases within the tRNA acceptor−TΨC minihelix.[9,13] Further work on the secondary structure of the cTAR[14−17] and the primer binding site (PBS) stem loop (SL)[18] again demonstrated the ability of NC to interfere with base stacking but without completely opening the helix. Known RNA annealing activities of NC include promoting dimerization of the RNA genome and positioning of the tRNA primer on the PBS. However, these processes occur in immature virions prior to the release of NC from the precursor Gag polyprotein,[2,19] suggesting that it is the uncleaved Gag that may be the effector molecule. Similarly, the highly specific capture of the viral genomic RNA from the vast excess of similar cellular RNAs must occur before budding and processing of the polyprotein, again with Gag necessarily being the active moiety involved. Gag has been less studied as an RNA chaperone than NC, although the NC domain of the Gag polyprotein is in part similar in structure to the free NC protein.[20,21]

Gag has also been shown to have superior RNA chaperoning ability compared to that of NC in strand transfer assays during

Received: December 7, 2011
Revised: March 5, 2012
Published: March 26, 2012

   dx.doi.org/10.1021/bi2017969 | *Biochemistry* 2012, 51, 3162−3169

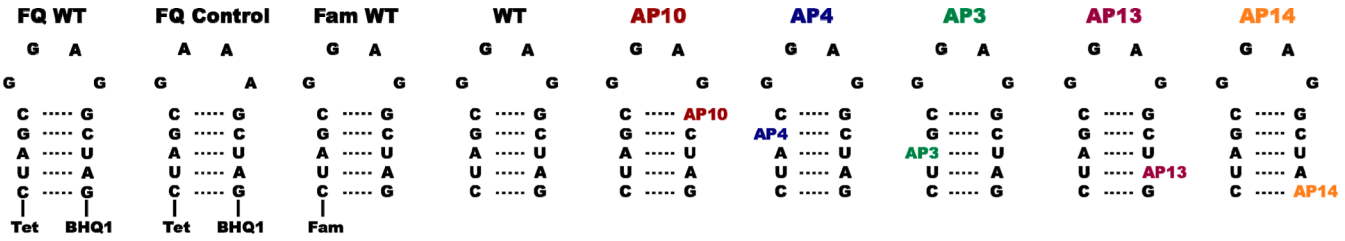


Figure 1. RNA substrates used in this study.

reverse transcription.[5] Recent evidence has emerged that the matrix domain of Gag may have important functional RNA binding activity;[22,23] however, while the matrix domain may have supplementary functions in viral assembly,[7] it is the NC region that is overwhelmingly dominant in RNA capture.[5] Despite this, far less is known about the ability of the Gag polyprotein to affect nucleic acid structure and in particular Gag's ability, compared to that of NC alone, to unwind an RNA helix.

SL3 is a key nucleic acid structure found within the Ψ-packaging domain in the 5′UTR of the HIV-1 RNA genome and plays a pivotal role in the packaging of genomic RNA. SLs within the Ψ-packaging domain are believed to provide Gag with an initial binding site for the genomic RNA, which is used to distinguish the genomic RNA from other cellular and viral RNAs present within the host cell.[2] Binding of Gag to the Ψ-packaging domain has been suggested to unwind the SL3 helix in vitro[4] and is thought to lead to the structural rearrangement of the RNA possibly exposing further high-affinity Gag binding sites. Structural rearrangement of the RNA, caused by Gag binding within the Ψ-packaging domain, is thus an essential step in the life cycle of HIV-1, and the ability of Gag to destabilize the SLs within the Ψ-packaging domain must play an important role.

To explore differences in the abilities of NC and Gag to unfold RNA structures, we undertook a systematic study using the 2-aminopurine-modified RNA substrates based on the 14-nucleotide GGAG tetraloop SL3. 2-Aminopurine has the ability to report upon its local environment while itself causing minimal structural distortion and as such is a robust and useful structural purine nucleobase substitute, which has been used extensively for such studies.[24−29] We replaced individual A or G nucleotides from each of the five base pairs within the stem of SL3 to create a series of five 2-aminopurine-substituted RNA substrates, **AP10**, **AP4**, **AP3**, **AP13**, and **AP14** (Figure 1). These RNA substrates were biophysically characterized and shown to form SLs comparable to that of wild-type SL3. Binding studies of each of the RNA substrates revealed differences in the ability of NC and Gag to affect the nucleic acid structure of the stem of SL3. Furthermore, the addition of Gag and NC to a molecular beacon designed using SL3 [by attachment of a 5′ Tet fluorophore and 3′ black hole quencher 1 (BHQ1) to the SL3 RNA] showed that Gag was able to significantly increase the fluorescence intensity of the molecular beacon when compared to NC only. The data from both the 2-aminopurine and molecular beacon studies suggest that Gag has a significantly greater ability to affect NA structure than NC does and confirm the Ψ helix unwinding ability of Gag in the packaging process.

MATERIALS AND METHODS

Synthetic RNA Oligonucleotides. All RNA oligonucleotides were synthesized and purified by IBA GmbH. The sequences of the wild-type, labeled, and 2-aminopurine-modified RNAs were as follows (Figure 1): **WT**, CUA GCG GAG GCU AG; **FQ WT**, Tet-CUA GCG GAG GCU AG-BHQ1; **FQ Control**, Tet-CUA GCG AAA GCU AG-BHQ1; **FAM WT**, FAM-CUA GCG GAG GCU AG; **AP10**, CUA GCG GAG XCU AG; **AP4**, CUA XCG GAG GCU AG; **AP3**, CUX GCG GAG GCU AG; **AP13**, CUA GCG GAG GCU XG; **AP14**, CUA GCG GAG GCU AX (where X denotes a 2-aminopurine-modified nucleotide). Stock solutions (100 μM) of all RNA oligonucleotides were made by resuspending the RNAs in molecular biology grade water and quantified by A_{260} at 95 °C, using ε_{260} values of 139.7 (**WT**, **FAM WT**, and **FQ WT**), 143.5 (**FQ Control**) 118.87 (**AP10**), 115.57 (**AP4**), 112.17 (**AP3**), 112.17 (**AP13**), and 121.37 $cm^2\ \mu mol^{-1}$ (**AP14**) as provided by the manufacturers, before the RNAs were aliquoted and stored at −80 °C.

Expression and Purification of the GagΔP6 Protein. GagΔP6 was expressed in *Escherichia coli* BL21DE3 pLys S cells with a GST tag and a Prescission protease cleavage site by subcloning the ECoRI- and SalI-digested insert from pGexGagΔP6[30] into EcoRI- and SalI-digested pGex6p-1 (GE). The cells were induced with 0.5 mM IPTG at 30 °C after growth at 37 °C to an OD_{600} of 1 and harvested after 90 min. The cells were resuspended [50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 1 mM DTT, and 0.1 mM $ZnCl_2$] and treated with lysozyme (0.5 mg/mL) for 30 min, before sonication. Triton X-100 was added to the cell lysate (to a final concentration of 0.5%), and then the lysate was clarified by centrifugation (12000*g* for 20 min), with the clarified cell lysate incubated with glutathione beads for 60 min (all steps performed at 4 °C). The beads were collected by filtration and washed repeatedly with a high salt buffer (1 M NaCl) followed by a low salt (150 mM NaCl) buffer [50 mM Tris-HCl (pH 7.5), 1 mM DTT, and 0.1 mM $ZnCl_2$]. Cleavage of the GagΔP6 protein was achieved by overnight treatment with Prescission protease (GE) in 150 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM DTT, and 0.1 mM $ZnCl_2$ at 4 °C. The beads were washed with low salt buffer before elution of the GagΔP6 protein in high salt buffer. The purified GagΔP6 protein was rebuffered [25 mM NaOAc (pH 6.5), 50 mM NaCl, 1 mM DTT, and 0.1 mM $ZnCl_2$] by gel filtration and quantified via Coomassie Brilliant Blue before the protein was aliquoted and stored at −80 °C. Lyophilized NC protein was resuspended in the same buffer that was used in the purification of GagΔP6 [25 mM NaOAc (pH 6.5), 50 mM NaCl, 1 mM DTT, and 0.1 mM $ZnCl_2$] before the protein was aliquoted and stored at −80 °C.

Fluorescence Intensity and Fluorescence Polarization (FP) Measurements of RNA Titrations with NC and GagΔP6 Protein. All samples were freshly prepared prior to each experiment.

dx.doi.org/10.1021/bi2017969 | *Biochemistry* 2012, 51, 3162−3169

Aliquots of **FAM WT**, **AP10**, **AP4**, **AP3**, **AP13**, or **AP14** RNA and NC or GagΔP6 protein were allowed to defrost on ice prior to being used. All RNA oligonucleotides were heated at 95 °C for 5 min and snap-cooled on ice for a further 5 min, before the addition of buffer to give an RNA concentration of 1 μM [25 mM NaOAc (pH 6.5), 2 mM $MgCl_2$, and 0.1 mM $ZnCl_2$]. Five microliters of the RNA substrate was added to the wells of a 384-well microplate (low-volume flat bottom black NBS treated, Corning 3820) before 5 μL of NC or GagΔP6 protein was added to the wells containing the RNA at concentrations between 0 and 20 μM. The microplate was incubated for 45 min before measurements were taken on a Pherastar$^+$ plate reader (BMG LabTech) at 25 °C. For the steady-state measurement, a 320 nm excitation filter (to minimize protein fluorescence) and a 380 nm emission filter were used. For FP measurements, a 485 nm excitation filter and a 520 nm emission filter were used. Each measurement is the mean of three independent experiments, and errors are one standard deviation (SD). Experimental titration data were modeled using a nonlinear least-squares fit of a single-site, two-state binding model equation (eq 1).[31] Dissociation constants (K_d) and apparent dissociation constants ($K_{d'}$) were estimated from the total change in polarization or the total change in 2-aminopurine fluorescence intensity, $(I - I_0)/(I_f - I_0)$, plotted as a function of the total NC or GagΔP6 concentration.

$$y = \frac{[R_{tot} + P_{tot} + K_d - \sqrt{(R_{tot} + P_{tot} + K_d)^2 - 4(R_{tot}P_{tot})}\,]}{2R_{tot}} \quad (1)$$

where I is the measured fluorescence intensity/polarization, I_0 is the initial fluorescence intensity/polarization of the RNA substrate in the absence of protein, I_f is the fluorescence intensity/polarization of the RNA substrate when saturated with protein, R_{tot} is the total RNA concentration, P_{tot} is the total protein concentration, and K_d is the dissociation constant (or $K_{d'}$ as the apparent dissociation constant).

Because of the high NC concentrations required to saturate the system when using **AP3**, **AP13**, and **AP14**, the assumption was made that I_f during NC titration would be the same as I_f during GagΔP6 titration. This assumption was based on the very similar experimental I_f gained for both **AP10** and **AP4** SLs, from saturation with NC and GagΔP6 titrations (Figures 6−10 of the Supporting Information).

Fluorescence Intensities of the Molecular Beacons FQ WT and FQ Control with Titrations of NC and GagΔP6 Protein. These experiments were essentially conducted as described above. Five microliters of the snap-cooled molecular beacon **FQ WT** or **FQ Control** at a concentration of 0.2 μM was added to the wells of a 384-well plate followed by addition of 5 μL of either the protein buffer, NC, or GagΔP6 (0−20 μM). The microplate was incubated for 45 min before measurements were taken at 25 °C using a Pherastar$^+$ plate reader (BMG LabTech) with a 510 nm excitation filter and a 540 nm emission filter. The **FQ WT** and **FQ Control** RNA in their native and chemically denatured (9 M urea) states were corrected for screening and buffer effects and normalized as 0 and 100% intensity, respectively; the signal gained from the unfolding of **FQ WT** and **FQ Control** by GagΔP6 or NC is described in comparison to these native and denatured states. Each measurement is the mean of three independent experiments, and errors are one SD and the data modeled with eq 1 as described above.

RESULTS

Design and Structural Validation of 2-Aminopurine-Modified RNA Oligonucleotides AP10, AP4, AP3, AP13, and AP14. The 14-nucleotide RNA SL used in this study is the native SL3 found in the Ψ-packaging domain of HIV-1. SL3 consists of a 5 bp stem with a four-base GGAG loop and as such is part of the *GNAR* family of tetraloops.[4] To study the interactions of both NC and GagΔP6 proteins on the overall structure of SL3, we systematically incorporated a 2-aminopurine-modified nucleotide at each of the 5 bp comprising the stem of SL3 (Figure 1). This gave a series of five 2-aminopurine-modified RNAs, which are capable of giving single-nucleotide resolution of the binding of NC and GagΔP6. Because of the differences in hydrogen bonding and sterics of the 2-AP:C and 2-AP:U base pairs compared to those of the native G:C and A:U base pairs, it was important to biophysically characterize the 2-aminopurine-modified RNA substrates compared to the **WT** RNA (Figures 2−5 of the Supporting Information)

The melting temperature (T_m) of **WT** (67.9 ± 0.5 °C) SL3 RNA indicates the formation of a highly thermally stable SL (Figure 2 of the Supporting Information). Furthermore, a magnesium titration of **WT** SL3, measured by CD (Figure 4 of the Supporting Information), indicates that very little structural change is observed over the range of magnesium concentrations measured (0−10 μM). The high thermal and structural stability of SL3 fits well with previous studies of RNA tetraloops.[32]

2-AP modification to nucleotides at positions 10 and 4 created the modified SLs **AP10** (T_m = 53.5 ± 0.8 °C) and **AP4** (T_m = 43.2 ± 0.4 °C). The large differences in the melting temperatures (ΔT_m) of **AP10** (ΔT_m = 14.4 °C) and **AP4** (ΔT_m = 24.7 °C) highlight the importance of the closing G-C base pair in the thermal stability of RNA tetraloops.[33−36] Modifications at positions 3, 13, and 14 gave SLs **AP3** (T_m = 63.2 ± 0.4 °C), **AP13** (T_m = 64.3 ± 0.8 °C), and **AP14** (T_m = 62.1 ± 0.8 °C). The relatively small differences in ΔT_m values for SLs **AP3** (ΔT_m = 4.7 °C), **AP13** (ΔT_m = 3.6 °C), and **AP14** (ΔT_m = 5.8 °C) indicate that modifications toward the lower part of the SL yield little change in their overall thermal stability. The thermal denaturation data show that at 25 °C the majority of the RNA population is in the folded SL structure, which applies to all of the 2-aminopurine-modified SLs. This is further supported by the CD spectra obtained at 25 °C, where all the 2-aminopurine-modified SLs suggest a structure closely related to that of the **WT** RNA (Figure 4 of the Supporting Information).

Binding of GagΔP6 and NC to Wild-Type SL3 Measured by Fluorescence Polarization. Each 2-aminopurine-modified RNA SL can provide detailed insight into the effects that binding of GagΔP6 and NC has on the local structure of SL3, while together the 2-aminopurine series can illustrate how the binding of the viral proteins affects the global structure of SL3. The 2-aminopurine modifications provide only apparent dissociation constants ($K_{d'}$) relating to each 2-aminopurine nucleotide. Dissociation constants for GagΔP6 and NC with SL3 as a whole were determined by measurement of the changes in the fluorescence polarization (FP) as protein was titrated against a 5′-FAM-labeled SL3 (**FAM WT**).

The determination of the K_d values for both GagΔP6 (0.086 ± 0.002 μM) and NC (0.117 ± 0.003 μM) was important

dx.doi.org/10.1021/bi2017969 | *Biochemistry* 2012, 51, 3162−3169

Table 1. Relative Melting Temperatures, Quantum Yields, and Dissociation Constants of the 2-AP RNA Substrates Used in This Study

	UV thermal denaturation				quantum yield						apparent dissociation constant ($K_{d'}$)[h]			
					folded[d]		unfolded[e]				NC		GagΔP6	
substrate	T_m(°C)[a]	1 SD	ΔT_m (°C)[b]	substrate	Φ[c]	1 SD	Φ	1 SD	ΔΦ[f]	substrate	$K_{d'}$ (μM)	1 SD	$K_{d'}$ (μM)	1 SD[g]
WT	67.9	0.4	–	2-AP	0.680	0.680	0.690	0.002	–	**Fam-WT**	0.117[i]	0.003	0.086[i]	0.002
AP10	53.5	0.5	14.4	**AP10**	0.190	0.006	0.117	0.007	0.073	**AP10**	0.10	0.04	0.21	0.03
AP4	43.2	0.4	24.7	**AP4**	0.008	0.0004	0.080	0.003	0.072	**AP4**	0.23	0.04	0.11	0.04
AP3	63.2	0.4	4.7	**AP3**	0.002	0.0002	0.020	0.001	0.018	**AP3**	1.4	0.2	0.25	0.05
AP13	64.2	0.8	3.7	**AP13**	0.004	0.0002	0.038	0.001	0.034	**AP13**	2.3	0.2	0.16	0.04
AP14	62.1	0.9	5.8	**AP14**	0.039	0.004	0.187	0.002	0.148	**AP14**	5.0	0.2	0.13	0.08

[a]Melting temperatures were determined in 25 mM NaOAc (pH 6.5), 25 mM NaCl, and 0.1 mM $MgCl_2$ at an RNA substrate concentration of 5 μM. [b]Difference in melting temperature compared to that of wild-type SL3. [c]Relative quantum yields were determined using 2-aminopurine (Φ = 0.68) under folded and unfolded conditions at an RNA substrate concentration of 0.5 μM. [d]Folded conditions included 25 mM NaOAc (pH 6.5), 25 mM NaCl, and 0.1 mM $MgCl_2$. [e]Unfolded conditions included 25 mM NaOAc (pH 6.5), 25 mM NaCl, and 9 M urea. [f]Difference in quantum yield under folded and unfolded conditions. [g]Each measurement is the mean of three independent experiments, and errors are one SD. [h]Apparent dissociation constant ($K_{d'}$) measured by 2-aminopurine intensities. [i]Dissociation constant (K_d) measured by FP.

(Table 1 and Figure 2A) as it rules out any differences seen from the $K_{d'}$ values determined when using the 2-aminopurine-modified RNA substrates, being an effect from the difference in the binding affinities of the Gag and NC proteins.

Binding of GagΔP6 and NC to the Series of 2-Aminopurine-Modified RNA Stem Loops. The $K_{d'}$ value for each 2-aminopurine-modified nucleotide can be combined to determine whether the effects of GagΔP6 and NC binding encountered at the top of the stem are translated equally through the 5 bp to the bottom of the stem and therefore whether the two proteins affect the SL in a similar fashion at comparable concentrations.

If one starts with **AP10** and **AP4** where the 2-aminopurine modifications are positioned at the top of the stem, the $K_{d'}$ values for NC (for **AP10**, $K_{d'}$ = 0.10 ± 0.04 μM, and for **AP4**, $K_{d'}$ = 0.23 ± 0.04 μM) and GagΔP6 (for **AP10**, $K_{d'}$ = 0.21 ± 0.02 μM, and for **AP4**, $K_{d'}$ = 0.11 ± 0.04 μM) are not only similar to each other but also similar to the K_d values determined by FP for the binding of NC (K_d = 0.117 ± 0.003 μM) and GagΔP6 (K_d = 0.086 ± 0.002 μM) to **FAM WT**. Moving down the stem to the third base pair below the GGAG tetraloop reveals a difference between the $K_{d'}$ values for NC ($K_{d'}$ = 1.4 ± 0.2 μM) and GagΔP6 ($K_{d'}$ = 0.25 ± 0.04 μM) for the **AP3** RNA substrate. This difference in $K_{d'}$ values between NC and GagΔP6 continues to increase at the fourth and fifth base pair below the GGAG tetraloop, with RNA substrates **AP13** (for NC, $K_{d'}$ = 2.3 ± 0.2 μM, and for GagΔP6, $K_{d'}$ = 0.18 ± 0.04 μM) and **AP14** (for NC, $K_{d'}$ = 5.0 ± 0.2 μM, and for GagΔP6, $K_{d'}$ = 0.13 ± 0.6 μM), respectively (Table 1 and Figures 6–10 of the Supporting Information).

Binding of GagΔP6 and NC to the Molecular Beacons FQ WT and FQ Control. To further describe the unwinding ability of GagΔP6 compared to that of NC, an assay involving an RNA molecular beacon (**FQ WT**) was established. **FQ WT** (T_m = 72.2 ± 0.6 °C) was titrated against either GagΔP6 or NC, in the same manner described for the 2-aminopurine-modified RNA substrates. Addition of NC to **FQ WT** (K_d = 0.33 ± 0.06 μM) results in an increase in the magnitude of the signal of ~15% at 10 μM, whereas the unwinding of **FQ WT** (K_d = 0.15 ± 0.02 μM) by GagΔP6 leads to an increase in the magnitude of the signal of ~60% at 10 μM, when compared to that of chemically denatured **FQ WT** (Figure 3B).

The ability of NC to push a structured nucleic acid population toward a more unfolded or folded state has been described previously, and the slight increase in the fluorescence from the **FQ WT** RNA is most probably in response to this shift in the population.[16] Addition of the GagΔP6 protein results in a significant change in the fluorescence intensity and further supports the idea that GagΔP6 has better SL destabilizing and unwinding ability than NC.

DISCUSSION

The biophysical characterization of the **WT** RNA shows that the GGAG tetraloop forms a highly stable hairpin structure.

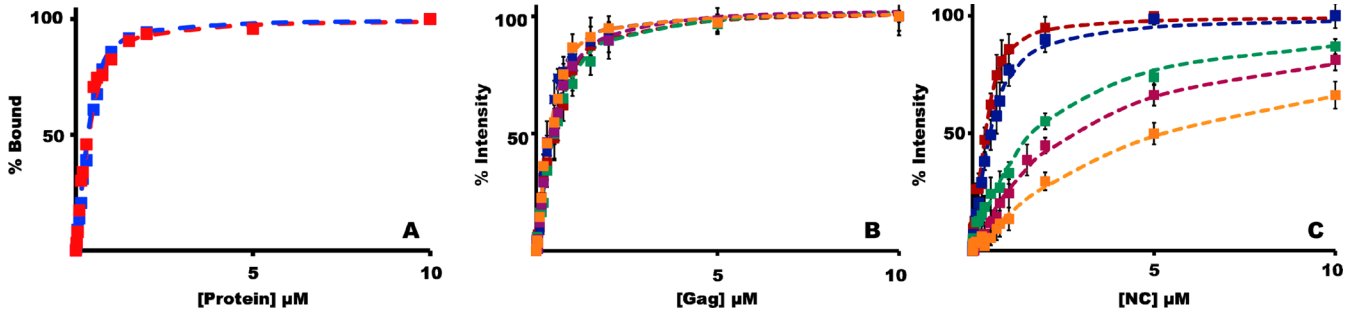


Figure 2. Titration curves of **Fam WT** and 2-aminopurine-modified RNA SLs **AP10** (dark red, top), **AP4** (dark blue), **AP3** (green), **AP13** (dark red, bottom), and **AP14** (orange) with GagΔP6 and NC. The changes in (A) fluorescence polarization [GagΔP6 (blue) and NC (red)] or (B and C) 2-aminopurine fluorescence intensity, $(I - I_0)/(I_f - I_0)$, were plotted as a function of the total NC or GagΔP6 concentration with an RNA substrate concentration of 0.5 μM. The dashed lines in panels A–C of Figure 3 correspond to the fits of the experimental points with eq 1.

dx.doi.org/10.1021/bi2017969 | *Biochemistry* 2012, 51, 3162–3169

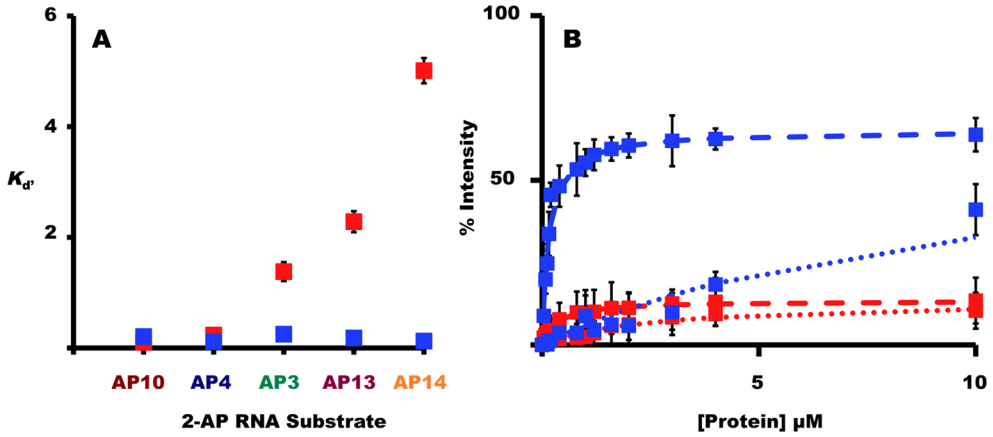


Figure 3. (A) Plot of apparent dissociation constant ($K_{d'}$) for GagΔP6 (blue) and NC (red) vs 2-aminopurine positional modification within the RNA substrates. Starting with **AP10** as the closing base pair to **AP14** as the terminal base pair. (B) Plot of the change in fluorescence intensity of **FQ WT** (———) and **FQ Control** (···) upon addition of GagΔP6 (blue) or NC (red) with an RNA substrate concentration of 0.1 μM. Fluorescence intensities, $(I - I_0)/(I_f - I_0)$, were plotted as a function of the total NC or GagΔP6 concentration. The dashed lines in panel B correspond to the fits of the experimental points to eq 1. Because of the high protein concentration that would be required to saturate the FQ control RNA, the dotted lines are illustrative, not actual fits of the experimental data.

The high thermal stability of SL3 has been reported previously by Borer et al. (T_m < 70 °C)[37] and Scholes et al. (T_m = 86.4 °C with Mg^{2+}, and T_m = 76.8 °C without Mg^{2+}),[38] with both groups using a 3 bp extended SL3.

The characterization of the modified RNAs demonstrated that a 2-aminopurine inserted at or just below the closing C-G base pairs has a stronger effect on the thermal stability of the SL than modifications made further down the stem. The ΔT_m values for the series of 2-aminopurine-modified SLs ranged from 3.6 to 24.7 °C, when compared to that of **WT** SL3 (Figure 2 of the Supporting Information). All the 2-aminopurine RNAs showed T_m values of >40 °C. Furthermore, CD spectra showed an A-form RNA structure for all the 2-aminopurine-modified SLs at 25 °C (Figure 3 of the Supporting Information), indicating that the 2-aminopurine series exhibit a SL structure that closely resembles the structure of **WT** SL3 (Figure 3 of the Supporting Information).

The quantum yield determinations for the 2-aminopurine RNA substrates in folded and unfolded states again support a SL structure (Figure 5 of the Supporting Information). **AP10** like **AP3**, **AP4**, and **AP13** has two neighboring nucleotides, yet **AP10** shows a very strong emission spectrum compared to those of **AP3**, **AP4**, and **AP13**. The structure of the GGAG tetraloop, determined by NMR spectroscopic studies,[4] shows that the G nucleobase at position 9 of the tetraloop is oriented away from the G nucleobase at position 10. Therefore, the limited base stacking of the G nucleobases at position 10 results in stronger emission spectra for **AP10** than for any of the other 2-aminopurine-modified SLs.

The binding of NC to SL3 has been extensively investigated using a variety of biophysical techniques,[37−45] whereas binding of Gag to SL3 has received far less attention. There are several published studies, using fluorescence anisotropy (FA), that compare the binding of NC and Gag to different RNAs. Musier-Forsyth et al. used a 37-nucleotide RNA minihelix^{Lys3} (for NC, K_d = 0.092 ± 0.009 μM, and for GagΔP6, K_d = 0.07 ± 0.02 μM);[23] Williams et al. used a 10-nucleotide $(UG)_5$ repeat (for NC, K_d = 0.012 ± 0.002 μM, and for GagΔP6, K_d = 0.019 ± 0.008 μM),[6] and Wu et al. used a 20-nucleotide (AGC UGC UUU UUG CCU GUA CU) sequence from the extreme 3′ terminus of U3 (for NC, K_d = 0.10 ± 0.02 μM, and for GagΔP6, K_d = 0.01 ± 0.02 μM).[5] The differences found in the binding of Gag and NC to NA have previously been discussed in the literature.[7]

As NC, and therefore the NC domain within GagΔP6, binds to the GGAG tetraloop at the top of the stem,[4,44] the similarity between the $K_{d'}$ values determined for **AP10** and **AP4** and the K_d values determined by FP above (with NC and GagΔP6) are consistent with the assumption that the local environments of the 2-aminopurine nucleotide, in **AP10** and **AP4**, are similar when bound to the free NC or to the NC domain within the GagΔP6 protein.

Further down the stem RNA substrates **AP3**, **AP13**, and **AP14**, the similarities among the $K_{d'}$ values determined for GagΔP6 and NC start to diverge with the distance of the base pair from the GGAG tetraloop (Figure 3A). The $K_{d'}$ values for GagΔP6 with **AP3**, **AP13**, and **AP14** all remain comparable with each other as well as with the $K_{d'}$ values determined for **AP10** and **AP4**.

Given that the NC footprint is approximately six to eight nucleotides and assuming a 1:1 RNA:NC stoichiometry,[44] the nucleotides at the bottom of the stem (**AP3**, **AP13**, and **AP14**) are less involved in a direct interaction with the free NC. The intensity change to the **AP3**, **AP13**, and **AP14** nucleotides may be more of a response to the destabilization of the base pairs caused by the accommodation of NC on the top of the loop or due to a second, nonspecific binding of NC further down the SL.

During the titrations of NC and GagΔP6 with **FQ WT**, the data from both titrations show little change in signal intensity at higher protein concentrations (Figure 3B). An expectation of a second nonspecific binding of NC to **FQ WT** might be the increased level of destabilization of the stem loop and therefore an increase in the magnitude of the signal measured as seen with the 2-aminopurine substrates mentioned above. The fact that with the **FQ WT** substrate there is very little signal change at higher NC concentrations suggests that a second NC binding is not occurring, but also the possibility that further opening of the stem loop does not occur during the second NC binding or that the fluorophore−quencher system interferes with this nonspecific binding.

dx.doi.org/10.1021/bi2017969 | *Biochemistry* 2012, 51, 3162−3169

The destabilization of nucleic acid structure by NC has been demonstrated previously by time-resolved fluorescence studies on a molecular beacon based on the DNA TAR SL; the population of the DNA TAR SL moved toward a more unfolded or folded state upon addition of NC.[14] To ensure the change in signal observed for the NC is not due to an artifact of the system used in this study, a second molecular beacon, **FQ Control** (T_m = 75.1 ± 0.3 °C), was employed in which the stem was kept identical to **FQ WT** but the loop was changed from a GGAG to a GAAA tetraloop. Previous studies determined that the SL3 GAAA mutation removed the high-affinity binding of NC to the tetraloop.[46] The titration of **FQ Control** with NC showed that nonspecific binding of the protein to the SL is possible (Figure 3B). We would argue that because of the size of the RNA chosen in this study, it is improbable that two NC proteins could fit onto the 14-nucleotide stem.

However, it is apparent that the initial, specific, and high-affinity interaction between NC and SL3 has little effect on the 2-aminopurine intensities (**AP3**, **AP13**, and **AP14**) toward the bottom of the stem, which is in agreement with the NMR solution structures of NC bound to an extended SL3. The NMR structure shows a single binding of NC and shows that this single NC is not enough to unwind the base pairs at the bottom of the stem.[44]

The footprint of GagΔP6 has not been established but is unlikely to be smaller than that of NC,[7] meaning the nucleotides at the bottom of the stem (for **AP3**, **AP13**, and **AP14**) would be more directly involved in the binding of GagΔP6. Our results support this, as all five 2-aminopurine-modified SLs as well as the molecular beacon **FQ WT** show consistent $K_{d'}$ values with GagΔP6. Again, the assumption is for a 1:1 binding of GagΔP6 to SL3, but there is the possibility that Gag exists in a multimeric state before the introduction of the RNA substrate or the RNA substrate acts as a template for the formation of higher-order structure. Rein et al. showed that GagΔP6 does dimerize in solution with a K_d of 5.5 μM,[47] and that aggregation of GagΔP6 is achieved by a nucleic acid template.[48]

The binding of more than one GagΔP6 on or near the RNA substrate would create a higher local concentration of the NC domain around the RNA substrate. To estimate stoichiometries of binding of GagΔP6 to the 2-AP RNA substrates, the data were modeled according to the method by Doudna et al.[49] Models for 1:1 and 1:2 RNA:GagΔP6 stoichiometries were also plotted as a comparison and illustrate that the data gained from the titration of GagΔP6 into the 2-AP substrates are best described by a 1:1 stoichiometry (Figures 6−10 of the Supporting Information). Furthermore, all the $K_{d'}$ values reported by the 2-aminopurine intensities are at least 20-fold below the K_d for Gag dimerization (Table 1).[47]

Gag has been shown to have a NA binding region in the matrix domain;[23] again the short SL used in our experiments is unlikely to be able to bind both NC and matrix regions on Gag simultaneously, and the dominant effect of NC means that effectively our system is studying NC domain binding alone. Further work to include larger RNAs and Gag mutants to further clarify the points made above and to provide further information about a possible mechanism for the unwinding of nucleic acid structure by Gag is ongoing.

For these reasons and because of the evidence of the molecular beacon experiments, we believe that the results presented here indicate that a much greater destabilizing effect is introduced through the stem to accommodate the binding of a single GagΔP6 to the top of the tetraloop, compared to NC.

CONCLUSION

The implication of a greater destabilizing effect by GagΔP6 on SL3 is that GagΔP6 shifts the population further toward a more unfolded state, suggesting that GagΔP6 is more able to alter NA structure than NC. The data are consistent with a model in which the Gag protein has the ability not only to bind to ssRNA, which is a function well established for the NC protein, but also to unwind the RNA helices to which it binds during the process of genome encapsidation. RNA structure in HIV is highly likely to be dynamic, with different regions adopting different structures during trafficking through the cell to accommodate the functional requirements of the RNA at different times (transactivation, nuclear export, translation and packaging, etc.). The viral genome is believed to be captured in the cytoplasm and transported to the plasma membrane chaperoned by relatively few Gag molecules[2] and at this stage as at many others is likely also complexed with cellular RNA binding proteins, many of which are still being identified. Structural switches also seem to be critical for this process, with accumulating evidence indicating that the viral genomic RNA can undergo quite striking changes in conformation to fulfill its function and these are undoubtedly triggered by RNA−protein interactions. The specific ability of Gag, rather than NC, to effect structural change after binding SL3 also prevents any prematurely cleaved NC within the cell from deflecting the RNA from its pathway to be packaged, as it will not trigger the same changes with their functional consequences.

ASSOCIATED CONTENT

Supporting Information

Additional materials and methods and further results and discussion of UV thermal melting, circular dichroism, quantum yield, and stoichiometry determination as well as individual plots for the titration of all 2-aminopurine-modified RNA substrates with NC and GagΔP6. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 OQQ, U.K. Telephone: +44 1223 336747. E-mail: amll1@medschl.cam.ac.uk.

Funding

This work was supported by a Medical Research Council Milstein Award (MRC 88493 to S.B. and A.M.L.L.) and the Biomedical Research Centre (Grant RG56162 to A.M.L.L.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Robert Gorelick for generously providing the NC protein.

ABBREVIATIONS

2-AP, 2-aminopurine; ΔΦ, difference in quantum yield; ΔT_m, difference in melting temperature; Φ, quantum yield; A, adenine; BHQ, black hole quencher; C, cytidine; CD, circular dichroism; FAM, 5(6)-carboxyfluorescein; G, guanine; HIV-1, human immunodeficiency virus type 1; I, measured fluorescence intensity/polarization; I_f, final fluorescence intensity/polarization; I_0, initial fluorescence intensity/polarization; K_d, dissociation constant; $K_{d'}$, apparent dissociation constant; Tet, tetrachlorofluorescein;

T_m, melting temperature; NC, nucleocapsid; PBS, primer binding site; P_{tot}, total protein concentration; R_{tot}, total RNA concentration; U, uridine; UTR, untranslated region; UV, ultraviolet.

REFERENCES

(1) Gheysen, D., Jacobs, E., Deforesta, F., Thiriart, C., Francotte, M., Thines, D., and Dewilde, M. (1989) Assembly and Release of HIV-1 Precursor Pr55Gag Virus-Like Particles from Recombinant Baculovirus-Infected Insect Cells. *Cell 59*, 103−112.

(2) Lever, A. M. (2007) HIV-1 RNA packaging. *Adv. Pharmacol. 55*, 1−32.

(3) Rein, A., Datta, S. A., Jones, C. P., and Musier-Forsyth, K. (2011) Diverse interactions of retroviral Gag proteins with RNAs. *Trends Biochem. Sci. 36*, 373−380.

(4) Zeffman, A., Hassard, S., Varani, G., and Lever, A. (2000) The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the gag polyprotein. *J. Mol. Biol. 297*, 877−893.

(5) Wu, T. Y., Datta, S. A. K., Mitra, M., Gorelick, R. J., Rein, A., and Levin, J. G. (2010) Fundamental differences between the nucleic acid chaperone activities of HIV-1 nucleocapsid protein and Gag or Gag-derived proteins: Biological implications. *Virology 405*, 556−567.

(6) Cruceanu, M., Urbaneja, M. A., Hixson, C. V., Johnson, D. G., Datta, S. A., Fivash, M. J., Stephen, A. G., Fisher, R. J., Gorelick, R. J., Casas-Finet, J. R., Rein, A., Rouzina, I., and Williams, M. C. (2006) Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res. 34*, 593−605.

(7) Rein, A. (2010) Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biol. 7*, 700−705.

(8) Levin, J. G., Guo, J. H., Rouzina, I., and Musier-Forsyth, K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. 80*, 217−286.

(9) Hargittai, M. R. S., Mangla, A. T., Gorelick, R. J., and Musier-Forsyth, K. (2001) HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing. *J. Mol. Biol. 312*, 985−997.

(10) Khan, R., and Giedroc, D. P. (1992) Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. *J. Biol. Chem. 267*, 6689−6695.

(11) Gregoire, C. J., Gautheret, D., and Loret, E. P. (1997) No tRNA3Lys unwinding in a complex with HIV NCp7. *J. Biol. Chem. 272*, 25143−25148.

(12) Khan, R., Chang, H. O., Kaluarachchi, K., and Giedroc, D. P. (1996) Interaction of retroviral nucleocapsid proteins with transfer RNAPhe: A lead ribozyme and ^{1}H NMR study. *Nucleic Acids Res. 24*, 3568−3575.

(13) Chan, B. D., Weidemaier, K., Yip, W. T., Barbara, P. F., and Musier-Forsyth, K. (1999) Intra-tRNA distance measurements for nucleocapsid protein-dependent tRNA unwinding during priming of HIV reverse transcription. *Proc. Natl. Acad. Sci. U.S.A. 96*, 459−464.

(14) Hong, M. K., Harbron, E. J., O'Connor, D. B., Guo, J. H., Barbara, P. F., Levin, J. G., and Musier-Forsyth, K. (2003) Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer. *J. Mol. Biol. 325*, 1−10.

(15) Bernacchi, S., Stoylov, S., Piemont, E., Ficheux, D., Roques, B. P., Darlix, J. L., and Mely, Y. (2002) HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol. Biol. 317*, 385−399.

(16) Cosa, G., Harbron, E. J., Zeng, Y. N., Liu, H. W., O'Connor, D. B., Eta-Hosokawa, C., Musier-Forsyth, K., and Barbara, P. F. (2004) Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys. J. 87*, 2759−2767.

(17) Cosa, G., Zeng, Y. N., Liu, H. W., Landes, C. F., Makarov, D. E., Musier-Forsyth, K., and Barbara, P. F. (2006) Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned opening of DNA hairpins. *J. Phys. Chem. B 110*, 2419−2426.

(18) Bourbigot, S., Ramalanjaona, N., Boudier, C., Salgado, G. F. J., Roques, B. P., Mely, Y., Bouaziz, S., and Morellet, N. (2008) How the HIV-1 Nucleocapsid Protein Binds and Destabilises the (−)Primer Binding Site During Reverse Transcription. *J. Mol. Biol. 383*, 1112−1128.

(19) Rein, A., Feng, Y. X., Campbell, S., Harvin, D., Ehresmann, B., and Ehresmann, C. (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: Possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol. 73*, 4251−4256.

(20) Summers, M. F., Newman, J. L., Butcher, E. W., Patel, D. T., and Mikhaylenko, Y. (2004) Flexibility in the P2 domain of the HIV-1 Gag polyprotein. *Protein Sci. 13*, 2101−2107.

(21) Datta, S. A., Curtis, J. E., Ratcliff, W., Clark, P. K., Crist, R. M., Lebowitz, J., Krueger, S., and Rein, A. (2007) Conformation of the HIV-1 Gag protein in solution. *J. Mol. Biol. 365*, 812−824.

(22) Barklis, E., Alfadhli, A., and Still, A. (2009) Analysis of Human Immunodeficiency Virus Type 1 Matrix Binding to Membranes and Nucleic Acids. *J. Virol. 83*, 12196−12203.

(23) Jones, C. P., Datta, S. A. K., Rein, A., Rouzina, I., and Musier-Forsyth, K. (2011) Matrix Domain Modulates HIV-1 Gag's Nucleic Acid Chaperone Activity via Inositol Phosphate Binding. *J. Virol. 85*, 1594−1603.

(24) Ward, D. C., Reich, E., and Stryer, L. (1969) Fluorescence Studies of Nucleotides and Polynucleotides. I. Formycin 2-Aminopurine Riboside 2,6-Diaminopurine Riboside and Their Derivatives. *J. Biol. Chem. 244*, 1228−1237.

(25) Jean, J. M., and Hall, K. B. (2002) 2-Aminopurine electronic structure and fluorescence properties in DNA. *Biochemistry 41*, 13152−13161.

(26) Bradrick, T. D., and Marino, J. P. (2004) Ligand-induced changes in 2-aminopurine fluorescence as a probe for small molecule binding to HIV-1 TAR RNA. *RNA 10*, 1459−1468.

(27) Hall, K. B., and Williams, J. (2004) Dynamics of the IRE RNA hairpin loop probed by 2-aminopurine fluorescence and stochastic dynamics simulations. *RNA 10*, 34−47.

(28) Dibrov, S. M., Johnston-Cox, H., Weng, Y. H., and Hermann, T. (2007) Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew. Chem., Int. Ed. 46*, 226−229.

(29) Godet, J., Ramalanjaona, N., Sharma, K. K., Richert, L., de Rocquigny, H., Darlix, J. L., Duportail, G., and Mely, Y. (2011) Specific implications of the HIV-1 nucleocapsid zinc fingers in the annealing of the primer binding site complementary sequences during the obligatory plus strand transfer. *Nucleic Acids Res. 39*, 6633−6645.

(30) Anderson, E. C., and Lever, A. M. (2006) Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation. *J. Virol. 80*, 10478−10486.

(31) Swillens, S. (1995) Interpretation of Binding Curves Obtained with High Receptor Concentrations: Practical Aid for Computer Analysis. *Mol. Pharmacol. 47*, 1197−1203.

(32) Sheehy, J. P., Davis, A. R., and Znosko, B. M. (2010) Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA 16*, 417−429.

(33) Antao, V. P., Lai, S. Y., and Tinoco, I. (1991) A Thermodynamic Study of Unusually Stable RNA and DNA Hairpins. *Nucleic Acids Res. 19*, 5901−5905.

(34) Blose, J. M., Proctor, D. J., Veeraraghavan, N., Misra, V. K., and Bevilacqua, P. C. (2009) Contribution of the Closing Base Pair to Exceptional Stability in RNA Tetraloops: Roles for Molecular Mimicry and Electrostatic Factors. *J. Am. Chem. Soc. 131*, 8474−8484.

(35) Williams, D. J., and Hall, K. B. (2000) Experimental and computational studies of the G[UUCG]C RNA tetraloop. *J. Mol. Biol. 297*, 1045−1061.

(36) Antao, V. P., and Tinoco, I. (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res. 20*, 819−824.

(37) Shubsda, M. F., Paoletti, A. C., Hudson, B. S., and Borer, P. N. (2002) Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry 41*, 5276−5282.
(38) Xi, X., Sun, Y., Karim, C. B., Grigoryants, V. M., and Scholes, C. P. (2008) HIV-1 nucleocapsid protein NCp7 and its RNA stem loop 3 partner: rotational dynamics of spin-labeled RNA stem loop 3. *Biochemistry 47*, 10099−10110.
(39) Shubsda, M. F., Kirk, C. A., Goodisman, J., and Dabrowiak, J. C. (2000) Binding of human immunodeficiency virus type 1 nucleocapsid protein to Ψ-RNA-SL3. *Biophys. Chem. 87*, 149−165.
(40) Clever, J., Sassetti, C., and Parslow, T. G. (1995) RNA Secondary Structure and Binding-Sites for Gag Gene-Products in the 5′-Packaging Signal of Human-Immunodeficiency-Virus Type-1. *J. Virol. 69*, 2101−2109.
(41) De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science 279*, 384−388.
(42) Maki, A. H., Ozarowski, A., Misra, A., Urbaneja, M. A., and Casas-Finet, J. R. (2001) Phosphorescence and optically detected magnetic resonance of HIV-1 nucleocapsid protein complexes with stem-loop sequences of the genomic Ψ-recognition element. *Biochemistry 40*, 1403−1412.
(43) Shubsda, M. F., Paoletti, A. C., Hudson, B. S., and Borer, P. N. (2002) Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry 41*, 5276−5282.
(44) De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science 279*, 384−388.
(45) Vuilleumier, C., Bombarda, E., Morellet, N., Gerard, D., Roques, B. P., and Mely, Y. (1999) Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: A fluorescence study. *Biochemistry 38*, 16816−16825.
(46) Paoletti, A. C., Shubsda, M. F., Hudson, B. S., and Borer, P. N. (2002) Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry 41*, 15423−15428.
(47) Datta, S. A. K., Zhao, Z., Clark, P. K., Tarasov, S., Alexandratos, J. N., Campbell, S. J., Kvaratskhelia, M., Lebowitz, J., and Rein, A. (2007) Interactions between HIV-1 Gag molecules in solution: An inositol phosphate-mediated switch. *J. Mol. Biol. 365*, 799−811.
(48) Campbell, S., and Rein, A. (1999) In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. *J. Virol. 73*, 2270−2279.
(49) Rambo, R. P., and Doudna, J. A. (2004) Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry 43*, 6486−6497.