

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology



journal homepage: http://ees.elsevier.com.jmb

Identification of a Minimal Region of the HIV-1 5'-Leader Required for RNA Dimerization, NC Binding, and Packaging

Xiao Heng¹[†], Siarhei Kharytonchyk²[†], Eric L. Garcia², Kun Lu¹, Sai Sachin Divakaruni¹, Courtney LaCotti¹, Kedy Edme¹, Alice Telesnitsky^{2*} and Michael F. Summers^{1*}

¹Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA ²Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109-0620, USA

Received 15 November 2011; received in revised form 13 January 2012; accepted 21 January 2012 Available online 27 January 2012

Edited by D. E. Draper

Keywords: retrovirus assembly; diploid genome packaging; RNA structure; ITC; nucleocapsid protein

Assembly of human immunodeficiency virus type 1 (HIV-1) particles is initiated in the cytoplasm by the formation of a ribonucleoprotein complex comprising the dimeric RNA genome and a small number of viral Gag polyproteins. Genomes are recognized by the nucleocapsid (NC) domains of Gag, which interact with packaging elements believed to be located primarily within the 5'-leader (5'-L) of the viral RNA. Recent studies revealed that the native 5'-L exists as an equilibrium of two conformers, one in which dimer-promoting residues and NC binding sites are sequestered and packaging is attenuated, and one in which these sites are exposed and packaging is promoted. To identify the elements within the dimeric 5'-L that are important for packaging, we generated HIV-1 5'-L RNAs containing mutations and deletions designed to eliminate substructures without perturbing the overall structure of the leader and examined effects of the mutations on RNA dimerization, NC binding, and packaging. Our findings identify a 159-residue RNA packaging signal that possesses dimerization and NC binding properties similar to those of the intact 5'-L and contains elements required for efficient RNA packaging.

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding authors. E-mail addresses: ateles@umich.edu; summers@hhmi.umbc.edu.

[†] X.H. and S.K. contributed equally to this work.

Present address: K. Lu, RNA Institute, State University of New York at Albany, 1400 Washington Avenue, Albany, NY 12222, USA.

Abbreviations used: HIV-1, human immunodeficiency virus type 1; NC, nucleocapsid; 5'-L, 5'-leader; ITC, isothermal titration calorimetry; RPA, RNase protection assay; RT, reverse transcriptase; VLP, virus-like particle; 2D, two-dimensional; NOESY, nuclear Overhauser enhancement spectroscopy.

Introduction

Like all retroviruses, the human immunodeficiency virus (HIV) specifically packages two copies of its unspliced RNA genome, both of which are utilized for strand-transfer-mediated recombination during reverse transcription.^{1,2} Although only one DNA allele is generated by this mechanism (and retroviruses are thus considered "pseudodiploid"),³ the production of recombinant proviruses from heterozygous virions appears to serve as a primary means for promoting genetic evolution under conditions of environmental and chemotherapeutic stresses.^{4–7} Genome selection is mediated by interactions between the nucleocapsid (NC) domains of a subset of the viral structural proteins (called Gag) and RNA elements located primarily within the 5'-leader (5'-L) region of the genome.^{8–18} Genetic studies indicate that genomes are selected for packaging as dimers,^{19,20} and there is now compelling evidence that dimerization and selection occur in the cytoplasm subsequent to nuclear export.^{21–23} HIV-1 genomes are trafficked to the plasma membrane by a small number of Gag proteins, where additional Gag proteins co-localize and virus assembly occurs.^{24,25} Subsequent to budding, the packaged RNA molecules exist as weakly formed, non-covalently linked dimers that form more stable dimers as the virus ages.^{26–31}

Agents that disrupt the structure of the NC domains of Gag and interfere with genome packaging have potent antiviral properties, 32-36 and a better understanding of the structural determinants of genome selection could facilitate the development of new therapeutic approaches for the treatment of AIDS. However, despite considerable effort, current understanding of the protein-RNA interactions³⁷ and mechanisms that promote and/or regulate retroviral genome selection remains limited (for reviews, see Refs. 8-17,38,39). For some retroviruses, including the Moloney murine leukemia virus and Rous sarcoma virus, relatively small regions of the 5'-L RNAs (fewer than 100 nucleotides) that are capable of directing the packaging of heterologous RNAs into virus-like particles (VLPs) have been identified, and this has facilitated high-resolution structural studies of the NC:RNA complexes responsible for RNA packaging.^{40–42} Although atomic-level structures have been determined for the HIV-1 NC protein bound to isolated hairpin and single-stranded RNAs,43-45 these studies did not explain how the virus discriminately packages its dimeric, unspliced genome. To date, no high-resolution structural information is available for a packaging-competent HIV-1 RNA.

Efforts to identify a minimal HIV-1 packaging signal have been complicated by a number of factors. First, the 5'-L contains elements that are important for functions that are unrelated to packaging but critical for replication (including transcriptional activation, primer binding during reverse transcription, splicing, and dimerization) (Fig. 1), and mutagenesis studies designed to test effects on packaging can adversely affect other functions. For example, mutagenesis studies originally suggested that the TAR hairpin is required for genome packaging, $^{48-50}$ but more recent studies indicate that effects originally attributed to a packaging defect were actually due to dominantnegative effects on viral gene expression.⁵¹ A second complication is that correct Rev-mediated RNA localization is a prerequisite to lentiviral RNA packaging, so that even RNAs that possess intact

packaging signals are inefficiently packaged with-A third issue is that HIV-1 VLPs can out Rev.⁵ efficiently assemble in the absence of their native genomes and, in the process, package similar amounts of cellular RNAs.⁵³⁻⁵⁷ Another complication relates to the fact that some previous packaging studies were conducted using transfection experiments in which the cellular concentration of mutant vector RNAs was high. At high cellular concentrations, the mutant RNAs can be packaged at high levels, as long as they do not have to compete with RNAs containing the wild-type 5'-L. In addition, efforts to determine the secondary structure of the HIV-1 5'-L using chemical probing, mutagenesis and biochemical approaches, and phylogenetic analyses⁵⁸⁻⁷⁰ have led to multiple structural proposals and mechanistic predictions, thus hindering the design of mutations to properly test structure/ function hypotheses.³

Recently, using an NMR approach for studying relatively large RNAs, we showed that the HIV-15'-L exists as an equilibrium of two conformationally distinct species, one in which residues spanning the gag start codon (G328-A356; hereafter called "AUG") form a hairpin and the GC-rich palindromic loop of the dimer promoting DIS hairpin is base paired with an upstream element (U5), and another in which AUG is base paired with U5 and the dimerpromoting loop of DIS is exposed.⁴⁶ U5:AUG formation, proposed originally on the basis of phylogenetic and biochemical analyses,^{64,71} promotes dimerization by displacing the DIS and simultaneously exposes high-affinity NC binding sites, thereby promoting the packaging of a dimeric genome.⁴⁶ Having identified the conformer of the 5'-L that directs packaging, we wished to explore the potential roles of specific elements within the leader on RNA packaging. Our present study identifies a 159-nucleotide segment of the 5'-L that is sufficient for RNA dimerization and NC binding in vitro and is required for efficient and selective RNA packaging.

Results

Construct design

A primary goal of these studies was to systematically remove elements from 5'-L (residues 1–356) without altering the folding of other structural elements, as assessed by the NC binding, dimerization, and NMR spectroscopic properties of the RNA. Under conditions of physiological-like ionic strength (PI buffer: 140 mM KCl, 10 mM NaCl, 1 mM MgCl₂, and 10 mM Tris–HCl, pH 7.0) and at an RNA concentration of ~1.0 μ M, 5'-L exists as an equilibrium mixture of two species: a monomeric



Fig. 1. (a) Representation of HIV-1 viral genome showing relative positions of splice sites and coding *versus* non-coding regions. (b) Secondary structure of the 5'-L conformer that promotes dimerization, NC binding, and genome packaging.⁴⁶ Long-range base-pairing interactions between U5 (blue) and AUG (green) help expose the dimer-promoting GC-rich loop of the DIS hairpin (cyan). Residues in pairing patterns that give rise to resolved, upfield-shifted adenosine C2-H signals are highlighted (colored boxes). The 3'-most residues of AUG (light green) are deleted in the 5'-L³⁴⁴ construct. (c) Mutations in AUG that affect base pairing with U5 influence the monomer–dimer equilibrium. Lane 1, native 5'-L exists as a mixture of monomer and dimer when incubating within PI buffer. Lane 2, substitution of 331-GAGAGAGGGUGC-GAGAGCGUCGGU-355 by 331-AACUAUACAAUCGGAGAGAGAUUUGA-355, which has been reported in our previous study known as 5'-L^{HP-AUG}, favors the AUG hairpin structure and exists mainly as a monomer. This mutation has been shown to inhibit packaging.⁴⁶ Lane 3, substitution of 333-UGGGUGC-339 by 333-CUCUAGA-339, which favors the AUG hairpin structure and been reported with packaging deficiency.⁴⁷ Lane 4, substitution of 337-UGGGUGCGAGAGAGGGUC-353 by 337-CGGGCACAAGAAAAA-353 (5'-L^{U5:,AUG}) favors the formation of U5:AUG base pairs and promotes dimerization and packaging.⁴⁶ Lane 5, 5'-L³⁴⁴, in which residues 345-356 were truncated from 5'-L. The AUG hairpin cannot form in 5'-L³⁴⁴ readily forms a dimer upon incubation in PI buffer. (e) Portion of the 2D NOESY spectrum obtained for a dimeric 5'-L³⁴⁴ RNA sample containing protons on the non-exchangeable adenosine C-2 and ribose (A^{2,R}), the guanosine ribose (G^R), and the cytosine ribose (C^R) carbons and deuterons at all other non-exchangeable aromatic and ribose carbons: A^{2,R}, G^R, C^R-[5'-L³⁴⁴]₂. NMR signal assignments were made based on comparisons with isolated oligo-RNAs corresponding to the TAR, lower-PBS, DIS, and Ψ -RNA hairpini

conformation in which the GC-rich palindrome of the DIS hairpin is sequestered by base pairing with U5 and AUG forms a hairpin, and a dimeric conformation in which AUG base pairs with U5 and the DIS is exposed (Fig. 1c) (dimer dissociation constant K_d = 0.9 ± 0.1 µM).⁴⁶ Mutations in AUG that favor U5:AUG base pairing promote dimerization and NC binding of 5'-L RNAs *in vitro* and the packaging of analogous vector RNAs into VLPs (Fig. 1c),⁴⁶ whereas the complete truncation of the AUG hairpin inhibits dimerization by preventing U5:AUG-dependent displacement of the DIS.⁴⁶ The present studies focused on 5'-L constructs that lacked the 3'-most residues of AUG (A345–A356 deleted; 5'-L³⁴⁴) (Fig. 1b).⁴⁶ This deletion allows U5: AUG formation while simultaneously precluding formation of the competitive AUG hairpin structure, thereby shifting the equilibrium in favor of the dimer (Fig. 1c and d).

Two-dimensional (2D) nuclear Overhauser enhancement spectroscopy (NOESY) spectra obtained the fully protonated [5'-L³⁴⁴]₂ RNA, exhibiting severe signal overlap, and were unassignable using standard methodologies.⁴⁶ However, a significant number of resolved adenosine C2-H signals were detectable in 2D NOESY spectra obtained for RNAs

in which all non-exchangeable aromatic protons except the C2-H of the adenosines were substituted by deuterium. For example, ¹H NMR signals with chemical shifts and NOE cross-peak patterns matching those observed for the isolated TAR, DIS, and Ψ site oligo-RNA hairpins were readily identified in 2D NOESY spectra obtained for a dimeric 5'-L³⁴⁴ sample that contained protons on the non-exchangeable adenosine C-2 and ribose (A^{2,R}), the guanosine ribose (G^R), and the cytosine ribose (C^R) carbons and deuterons at all other non-exchangeable aromatic and ribose carbons (A^{2,R},G^R,C^R-[5'-L³⁴⁴]₂, Fig. 1e) (superscript 8, 2 and R: protonated on the C-8, C-2, and ribose carbons, respectively). These outlier signals and their NOE interactions with H1' protons were monitored to assess potential effects of mutations and deletions on the structure of the RNA.

Additional leader RNAs that contained deletions of the TAR hairpin (5'-L^{344- Δ TAR}), the upper region of the PBS loop (5'-L^{344- Δ PBS</sub>), the TAR and upper PBS loop (5'-L^{344- Δ TAR- Δ PBS</sub>), and the TAR and Poly(A) hairpins and upper PBS loop (5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS) were prepared (Fig. 2a). RNA homogeneity was confirmed by denaturing polyacrylamide gel electrophoresis (PAGE, Fig. 2b). Base-pairing patterns of the TAR and Poly(A) hairpins and the lower stem of the PBS loop have been supported by several studies,^{64,72–76} including compensatory mutagenesis-based packaging experiments.⁷⁷ The PBS deletion construct was designed such that the entire upper region of the loop (residues A132–A216) was substituted by a GAGA}}} tetraloop (a member of the well-characterized GNRA family^{78–85}). Resolved A220 C2-H signals were observed for all constructs in partially deuterated samples (prepared as described above), including those containing both the native and truncated PBS loops (Fig. 2c), indicating that base pairing in the lower stem of the PBS loop is maintained in both the native and mutant 5'-L RNAs. 2D NOESY data obtained for the 5'-L³⁴⁴ mutants did not exhibit line broadening or additional NMR signals (other than those expected for the GAGA tetraloop) that would be indicative of alternative folding. In particular, the

Effects of TAR, Poly(A), and PBS deletions on dimerization and NC binding

assignments for these elements.

outlier adenosine C2-H signals of the most highly truncated construct, $5'-L^{344-\Delta TAR-\Delta Poly(A)-\Delta PBS}$, exhib-

ited ¹H NMR chemical shifts and 2D NOESY cross peaks similar to those observed for the wild-type 5'- L^{344} RNA (Fig. 2d). Signals assigned to the TAR,

Poly(A), and upper PBS loop structures were missing in spectra obtained for the corresponding deletion

mutants, supporting the ¹H NMR chemical shift

The 5'-L³⁴⁴, 5'-L^{344-ΔTAR}, 5'-L^{344-ΔPBS}, 5'-L^{344-ΔTAR-ΔPBS}, and 5'-L^{344-ΔTAR-ΔPoly(A)-ΔPBS} RNAs all exist predominantly as monomers at low ionic strength but form dimers upon incubation in PI buffer (Figs. 1d and 3a), indicating that the TAR, Poly(A), and upper PBS hairpin structures do not play a significant role



Fig. 2. (a) RNA constructs used in the present study: (1) 5'-L, (2) 5'-L³⁴⁴, (3) 5'-L^{344- Δ TAR</sub>, (4) 5'-L^{344- Δ PBS</sub>, (5) 5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS, and (6) 5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS. (b) Denaturing PAGE gels show sample purity and their relative electrophoretic migration. (c) Portions of 2D NOESY spectra showing the A220 C2-H peaks observed for (from left to right) A^{2,R}, G^R, C^R-labeled [5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS, for 5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS, and for a fully protonated oligoribonucleotide corresponding to the lower PBS hairpin (5'-ggCUCUGGUgagaGCCAGAGC; lower case, non-native), showing that the lower PBS stem structure is maintained in these 5'-L constructs. (d) Chemical shifts of the C2-H signals for the PBS lower stem, Ψ , and DIS hairpins are unaffected by removal of the TAR, Poly(A), and PBS upper loop [gray and black spectra correspond to RNAs labeled 2 and 6 in (a), respectively].}}}}}}



Fig. 3. Effects of TAR, Poly(A), and PBS deletions on NC binding. (a) Native gel electrophoresis data (2% high-gelling-temperature agarose gel containing Tris-borate buffer and ethidium bromide for RNA visualization) showing that $5'-L^{344-\Delta TAR}$, $5'-L^{344-\Delta TBS}$, $5'-L^{344-\Delta TBS}$, and $5'-L^{344-\Delta TAR-\Delta Poly(A)-\Delta PBS'}$ (labeled 3-6, respectively, as in Fig. 2) form dimers after incubation in PI buffer (140 mM KCl, 10 mM NaCl, 1 mM MgCl₂, and 10 mM Tris-HCl, pH 7.0), indicating that deletions of TAR, Poly(A), and PBS do not affect RNA dimerization. (b) ITC data obtained upon titration of NC into $5'-L^{344}$. The raw data are shown in the top panel, and the integrated heat changes are shown in the bottom panel. The first two calorimetric data points were deleted before nonlinear least-squares isotherm fitting using "Two Sets of Sites" mode (MicroCal Origin 5.0). The first set of binding parameters was utilized to characterize the highaffinity interactions between 5'-L and NC (Table 1). (c) ITC NC titration

data obtained for 5'-L^{344- Δ TAR</sub>, 5'-L^{344- Δ TAR</sub>, and 5'-L^{344- Δ TAR- Δ PBS</sub>, showing that these deletions have modest effects on highaffinity NC binding. (d) ITC NC titration data obtained for 5'-L^{344- Δ TAR- Δ POly(A)- Δ PBS. The post-high-affinity endothermic contributions to NC binding observed for 5'-L³⁴⁴ (and at diminished levels in the other 5'-L³⁴⁴ mutants) are nearly undetectable.}}}}

in dimerization. The potential roles of these structures in NC binding were assessed by isothermal titration calorimetry (ITC).⁴⁶ As observed previously in ITC studies of native and mutant forms of 5'-L, the binding isotherms obtained for 5'-L³⁴⁴ upon NC titration in PI buffer included an initial exothermic component followed by an endothermic component (Fig. 3b). The exothermic events are attributed to high-affinity NC binding since similar ITC profiles have been observed for oligo-RNAs that contain a single high-affinity NC binding site.^{43–45} The endothermic events occur during the later injections at progressively higher [NC]:[RNA] ratios and are attributed to weak NC–RNA interactions and NC- induced RNA unfolding.⁴⁶ 5'-L³⁴⁴ exists as a dimer under conditions of the ITC experiments, and nonlinear least-squares fitting using a two-event binding model indicates that $[5'-L^{344}]_2$ binds 15 ± 1 NC molecules per RNA strand with high affinity (K_d =17±9 nM) (Fig. 3b and Table 1). Both the number of high-affinity NC sites and the binding affinity are similar to those measured for 5'-L RNAs containing mutations in AUG that favor U5:AUG base pairing (5'-L^{U5:AUG}; 16±1 NC sites per strand; Table 1).⁴⁶

The ITC isotherms obtained upon titration of $5'-L^{344-\Delta TAR}$, $5'-L^{344-\Delta PBS}$, $5'-L^{344-\Delta TAR-\Delta PBS}$, and $5'-L^{344-\Delta TAR-\Delta Poly(A)-\Delta PBS}$ with NC were similar in

Table 1. Thermodynamic parameters from ITC measurements

	N_1	K_{d1} (nM)	ΔH_1 (kcal/mol)	$-T\Delta S_1$ (kcal/mol)	N_2	K_{d2} (mM)	ΔH_2 (kcal/mol)	$-T\Delta S_2$ (kcal/mol)
5'-L ³⁴⁴	15 ± 1	17±9	-12.5 ± 0.2	1.7 ± 0.5	31 ± 4	1.0 ± 0.5	-2.3 ± 9.0	-8.3 ± 8.6
$5'-L^{344-\Delta TAR}$	13 ± 1	17 ± 15	-14.0 ± 2.0	2.8 ± 2.9	30 ± 3	0.6 ± 0.4	1.5 ± 1.9	-12.6 ± 0.6
$5'-L^{344-\Delta PBS}$	11 ± 1	16 ± 15	-16.7 ± 1.8	5.7 ± 1.3	43 ± 9	4.2 ± 6.0	3.4 ± 0.3	-11.4 ± 1.0
$5'-L^{344-\Delta TAR-\Delta PBS}$	9±1	21 ± 15	-15.5 ± 1.9	4.7 ± 1.9	31 ± 3	0.8 ± 0.6	3.1 ± 0.2	-11.7 ± 0.7
$5'-L^{344-\Delta TAR-\Delta Poly(A)-\Delta PBS}$	8 ± 1	26 ± 39	-11.7 ± 6	4.4 ± 1.4		_	—	—

Two-site binding mode was utilized to fit the ITC isotherms from NC titrations. For those with no data (marked with an em dash), the magnitude of the endothermic component was too small for reliable calculation.

appearance to those obtained for 5'-L and 5'-L³⁴⁴ (Fig. 3b and c), except that these truncated RNAs bound fewer NC molecules with high affinity. Deletion of the TAR and PBS loops resulted in losses of 2 ± 1 and 4 ± 1 NC binding sites, respectively, and the simultaneous deletion of both the TAR and PBS loops eliminated 6 ± 1 NC sites (Table 1), indicating that NC binds independently and non-cooperatively to these hairpins. The isolated oligo-TAR RNA was previously shown to interact with NC,⁸⁶ and an oligo-RNA corresponding to the isolated TAR hairpin was found by ITC to bind a single NC molecule with high affinity $(K_d = 110 \pm 2 \text{ nM})$ (X.H. and M.S., unpublished results). Removal of the Poly(A) hairpin from 5'-L^{344- Δ TAR- Δ PBS} resulted in the loss of a single additional high-affinity NC site (Table 1). Thus, TAR, PBS, and Poly(A) hairpins appear to contribute approximately 2, 4, and 1 high-affinity NC binding sites, respectively. Interestingly, we also observed small but systematic reductions in the magnitude of the endothermic component of the ITC isotherms upon removal of TAR, Poly(A), and PBS elements (Fig. 3b and c), with isotherms for 5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS exhibiting only} a slight endothermic feature at the higher NC/RNA ratios (Fig. 3c). Thus, the NC-induced RNA unwinding that occurs after saturation of the high-affinity NC binding sites appears to diminish with decreasing size of the RNA.

Effects of TAR, Poly(A), and PBS deletions on RNA packaging

When HIV-1 particles bud from cells in the absence of genomic RNAs with packaging signals, these particles encapsidate "random" host mRNAs roughly in proportion to their intracellular concentrations.⁸⁷ In contrast, most such adventitiously packaged mRNAs are displaced when RNAs with authentic packaging signals are present, due to the specificity of HIV-1 for its own genome. Here, a packaging assay based on these observations was used to address whether or not the TAR, Poly(A), and PBS loop deletion RNAs studied above possessed information sufficient to mediate authentic HIV-1 RNA packaging. In these experiments, an RNA transcribed from the HIV-1-based helper function plasmid pCMV Δ R8.2 (herein called the Ψ^- helper),⁸⁸ which expresses all virion proteins except env but lacks all HIV-1 5'-L sequences upstream of the gag start codon, served as an internal control for the extent of packaging of an RNA that lacks all HIV-1 packaging elements. In the present studies, we adjusted transfection levels such that the test and Ψ^- helper RNAs were transcribed in cells at similar levels, thereby enabling a more quantitative assessment of relative packaging efficiencies (Fig. 4). RNA packaging levels were measured by using reverse transcriptase (RT)



Fig. 4. TAR, Poly(A), and PBS deletions do not affect RNA packaging. RNA packaging efficiency of 5'-L variants was monitored by RPA. (a) RNA samples obtained from the co-expression of a series of packaging test vectors with CMV Δ R8.2, an HIV-1 Ψ^- helper construct. Lanes 1–5, products of RNA samples harvested from media of cells transfected with the indicated constructs plus pCMV Δ R8-2 helper; lane 6, from cells transfected with pCMV Δ R8-2 alone; lane 7, from mocktransfected cells. Lanes 8-14 are RPA products of cellular samples from the same transfections used in lanes 1-7. P, undigested riboprobe; $\Delta pA = \Delta poly(A)$. The mobilities of size standards are indicated on the left; the mobilities of riboprobe fragments protected by transcripts from each of the test vectors and from the Ψ^- helper are indicated on the right. (b) RPA of viral samples using two different riboprobes specific to 7SL RNA or HIV-1 vector RNA (Ps, the two undigested riboprobes). Lanes 1-5, products of RNA samples harvested from media of cells transfected with the indicated constructs plus CMV Δ R8-2 helper; lane 6, from cells transfected with CMV Δ R8-2 alone; lane 7, from mock-transfected cells. The mobilities of riboprobe fragments protected by transcripts from each of the test vectors and the Ψ^- helper are indicated on the right as HIV RNA; the mobilities of riboprobe products protected by 7SL RNA are indicated as 7SL RNA.

activities (Fig. 4a) or viral 7SL RNA levels (Fig. 4b) as controls. 7SL RNA is a cellular transcript that is incorporated independent of genomic RNA and can be used as an internal control for quantitating packaging efficiencies.⁸⁹ Relative efficiencies were determined by comparing packaging levels of test RNAs that contained 5'-L mutations with those that contained the intact, wild-type leader and *gag* sequences. All test vectors contained Rev-responsive elements and were expressed in the presence of Rev, and thus Rev-dependent RNA localization, which is required for efficient packaging,⁵² was not a variable in these experiments.

As shown in the RNase protection assay (RPA) in Fig. 4a, when the Ψ^- helper was transfected into 293T cells in the absence of RNAs possessing HIV-1 packaging signals, its transcript was highly expressed in cells (lane 13) and was encapsidated into virions that were released from these cells with modest efficiency (lane 6). In contrast, whereas both transcripts were detected at similar levels within cells (lane 8), when CMV Δ R8.2 was co-transfected with an HIV-1 test vector possessing the native NL4-3 5'-L (Native), RNAs with native NL4-3 packaging elements were significantly enriched among virion RNAs, demonstrating that the leaderless Ψ^- helper transcripts could not compete for packaging (lane 1). TAR, Poly(A), and PBS loop deletion test RNAs also were co-expressed with the Ψ^- helper. Although these vectors lacked TAR regulatory elements, they were well expressed under the transfection conditions used, which were selected to ensure that the test vector and Ψ^- helper RNA levels were similar to one another within cells (lanes 9-12). However, similar to results obtained for the native NL4-3 leader, the Ψ^- helper RNAs were unable to compete efficiently with these test RNAs (lanes 2-5).

Similar results were obtained in RPAs that quantified vector RNA packaging by normalizing to the co-packaged host RNA 7SL, which serves as a surrogate for virion copy number⁹⁰ (Fig. 4b). Similar amounts of 7SL RNA-protected products were present in all viral samples from cells transfected with CMV Δ R8.2, whether alone or in the presence of test plasmids (lanes 1-6). As expected, HIV-1specific transcripts were robustly packaged in the sample from cells co-transfected with $CMV\Delta R8.2$ plus the parental 5'-L HIV-1 test vector (lane 1), and a significant drop in packaging was observed for the Ψ^- transcript produced when CMV Δ R8.2 was expressed alone (lane 6). In contrast to this $\Psi^$ transcript, 5'-L RNAs from which either the PBS loop or TAR was deleted (lanes 2-3), those that lacked both TAR and the PBS loop (lane 4), and the TAR, Poly(A), and PBS loop triple-deletion test RNA (lane 5), were all encapsidated at levels similar to those of native 5'-L RNA (lane 1). As quantified in Table 2, only small decreases in overall packaging were observed, with all of these deleted 5'-L RNAs displaying packaging >80% that observed for the native 5'-L.

Effects of gag deletion on packaging

To determine if elements in *gag* downstream of the AUG are important for packaging, we generated packaging vector variants that retained the 5'-L sequences studied by NMR above, but that lacked HIV-1 *gag*-derived sequences downstream of position +350 that were present in the native 5'-L vectors examined above. As shown in RPA results in Fig. 5, co-expression of either the original 5'-L or a truncated 5'-L³⁵⁰ vector lacking *gag* sequences with the CMV Δ R8.2 helper construct yielded indistinguishable RNA levels in cells (lanes 5 and 6) and identical strong packaging levels in virions (lanes 1 and 2; Table 2). These findings demonstrate that sequences in *gag* are not required for efficient packaging of HIV-1 5'-L containing RNAs into VLPs.

Effect of TAR deletion on 5'-L structure

The observation that deletion of TAR results in a moderate but reproducible reduction in RNA packaging, whereas deletion of the PBS loop does

Table 2. Packaging efficiencies of vector RNAs containing native and mutant HIV-1 5'-L sequences

Test RNA	%Packaging/RT ^a	%Packaging/ Ψ^- helper ^b	%Packaging/7SL ^c	
5'-L (Native)	100	100	100	
$5'-L-\Delta TAR$	41 ± 32	98 ± 4	86 ± 4	
$5'-L-\Delta PBS$	133 ± 77	100 ± 3	94 ± 4	
$5'$ -L- Δ TAR- Δ PBS	37 ± 20	90 ± 6	82 ± 1	
$5'$ -L- Δ TAR- Δ Poly(A)- Δ PBS	86 ± 59	100 ± 7	89 ± 11	
5'-L ³⁵⁰	101 ± 5	100 ± 1	ND	

Data reported as mean \pm standard deviation obtained from three independent experiments except data for 5'-L³⁵⁰, which are obtained from two independent experiments.

ND, not measured.

^a %Test RNA/virion based on RT activity to quantify virus production. Data normalized relative to test RNA containing the native 5'-L.

^b %Test RNA relative to the Ψ^- helper RNA in particles normalized relative to test RNA containing the native 5'-L.

^c %Test RNA/virion using 7SL packaging to quantify virus production. Data normalized relative to test RNA containing the native 5'-L.



Fig. 5. Effect of *gag* deletion on RNA packaging. RNA samples obtained from the co-expression of a series of native and 3' truncated packaging test vectors with CMV Δ R8.2. Lanes 1 and 2, RPA products of RNA samples harvested from media of cells transfected with the indicated constructs plus pCMV Δ R8-2 helper; lane 3, from cells transfected with pCMV Δ R8-2 alone; lane 4, from mock-transfected cells. Lanes 5–8 are RPA products of cellular samples from the same transfections used in lanes 1–4.

not affect packaging, is difficult to reconcile with our finding that the PBS upper loop contains a greater number of high-affinity NC binding sites. 2D NOESY spectra obtained for [5'-L^{344-ΔTAR}]2 appeared similar to spectra obtained for the parent $[5'-L^{344}]_2$ RNA (except for the missing TAR signals; data not shown), suggesting that the TAR deletion does not affect the overall fold of the dimeric RNA. To determine if the TAR deletion affects the folding of the monomeric form of the leader, we examined the dimerization properties and 2D NOESY spectra of leader RNAs in which AUG was fully truncated $(5'-L^{\Delta AUG})$. As indicated above, this RNA lacks the 5'-residues of AUG necessary to displace and expose the DIS hairpin, and like full-length 5'-L RNAs containing mutations that stabilize the AUG hairpin, 5'-L^{ΔAUG} exists predominantly as a monomer in PI buffer (Fig. 6a). Unexpectedly, deletion of the TAR hairpin from $5'-L^{\Delta AUG}$ ($5'-L^{\Delta AUG}$ - ΔTAR) strongly promoted dimerization in PI buffer (Fig. 6b). In addition, whereas the outlying adenosine C2-H signals in 2D NOESY spectra obtained for 5'-L $^{\Delta AUG}$ were similar in appearance to spectra obtained for $[5'-L^{344}]_2$ (Fig. 6b), the spectrum obtained for 5'-L^{$\Delta AUG-\Delta TAR$} lacked several expected signals, including the A268 C2-H signal associated with the properly folded form of the DIS hairpin (Fig. 6c). These findings indicate that removal of



Fig. 6. (a) Non-denaturing gel electrophoresis (2% agarose) results showing that $5'-L^{\Delta AUG}$ exists predominantly as a monomer, as expected, and that truncation of TAR (5'- $L^{\Delta AUG-\Delta TAR}$) strongly promotes dimerization. (b) [5'- L^{344}]₂ (green) *versus* $5'-L^{\Delta AUG}$ (black). (c) $5'-L^{\Delta AUG}$ (black) *versus* $5'-L^{\Delta AUG-\Delta TAR}$ (red). Stars highlight the expected loss of the TAR signals, and the arrow denotes the unexpected loss of the DIS A268-C2-H signal that was observed upon deletion of TAR. These data indicate that removal of TAR leads to misfolding of DIS hairpin and induces non-native dimerization of the monomeric form of the 5'-L.

TAR leads to non-native-like folding and dimerization of the 5'-L.

Discussion

The TAR, Poly(A), and PBS loops have each been implicated in at least some studies to play roles in genome dimerization, NC binding, and packaging. Isolated TAR oligo-RNA hairpins can form dimers in the presence of NC, and electron microscopy images of NC-incubated HIV-1 5'-RNA fragments (nucleotides 1–744) exhibit distinct circular structures that have been interpreted in terms of a TAR- mediated dimer interface.⁹¹ Some studies suggested that deletion of TAR promotes the packaging of monomeric RNAs, to an extent similar to that observed upon deletion of the DIS hairpin.92 On the other hand, truncation of the TAR hairpin from a 290-nucleotide 5'-fragment of the HIV-1 leader promoted, rather than inhibited, dimerization of this RNA. 93,94 The present studies show that 5'-L³⁴⁴, which spans the entire HIV-1 5'-untranslated region, forms a stable dimer even in the absence of the TAR hairpin. 2D NOESY spectra are relatively unperturbed (except for the loss of the TAR signals) upon removal of TAR, indicating that dimerization does not result from artificial misfolding of the dimeric leader. The fact that the A46 C2-H signals of TAR are relatively intense in both the monomeric and dimeric forms of 5'-L RNAs (Fig. 1e) suggests that the hairpin is relatively mobile and probably does not participate in long-range interactions of the dimeric 5'-L (although TAR could participate in intermolecular RNA-RNA interactions in the more stable dimer that is formed in virions during viral maturation^{92,95}).

Deletion of TAR was shown in one study to reduce packaging efficiency by 74% compared to the packaging of a parent RNA derived from the intact viral genome that encoded a stop codon within the capsid domain of *gag*.⁴⁹ Although TAR plays a critical role in transcriptional activation via its interactions with the viral Tat protein,96,97 these studies concluded that reductions in packaging were too great to be attributed to the relatively modest reductions in cellular RNA levels.⁴⁹ Mutations that destabilized the stem of TAR reduced packaging to 70% of wild-type levels (after controlling for reductions in cellular RNA levels)⁵⁰ and altered the ratio of spliced/unspliced viral RNAs that were packaged.⁹⁸ Interestingly, using a proviral expression system that is largely independent of Tat for transcriptional activation, Helga-Maria et al. found that missense mutations in the upper stem of TAR had little to no effect on packaging, but mutations at the base of the TAR stem could lead to significant reductions in packaging.⁴⁸ Compensatory mutations in the TAR stem restored packaging to near-wild-type levels, indicating that it is the structure of TAR, and not the TAR sequence, that is important for packaging.⁴⁸ A more recent study showed that the complete removal of TAR had little impact on replication as long as other changes were made to overcome the requirement for TARmediated transcriptional activation (packaging efficiencies were not reported).⁵¹ These studies also showed that there is evolutionary pressure to restore base pairing at the 5'-terminal end of the viral RNA⁵¹ and supported proposals that packaging defects induced by mutations in TAR result from misfolding of the 5'-L. Our findings are consistent with this latter hypothesis. Thus, deletion of TAR

did not significantly affect the NC binding properties of the 5'-L RNA and also did not significantly impair vector RNA packaging. The modest packaging defects that we observed upon deletion of TAR are most likely due to misfolding of the monomeric form of the 5'-L, which we believe leads to a reduction in the cellular levels of the packageable, dimeric form of the RNA.

The Poly(A) hairpin has also been proposed to play roles in both genome dimerization and packaging. Although early computer-modeling studies and electron microscopy imaging of the HIV-1 genome suggested that the palindromic 5'-AAGCUU-3' loop of Poly(A) might form a second dimer linkage structure,⁹⁹ subsequent mutagenesis studies indicated that the Poly(A) palindrome does not influence RNA dimerization.¹⁰⁰ Here, we showed that deletion of the TAR and Poly(A) hairpins does not affect the dimerization properties of 5'-L³⁴⁴, in support of the latter conclusion. Poly(A) was originally thought to be important for RNA packaging because removal of the hairpin reduced packaging to ~10% of wild-type levels.^{49,101} Mutations that disrupt the Poly(A) hairpin structure were also shown to inhibit genome packaging and viral replication, whereas compensatory mutations restored packaging to near-wild-type levels, indicating that it is the structure, rather than the specific sequence of the Poly(A) hairpin that is required for packaging.⁹⁸ However, more recent studies indicated that disruption of the Poly(A) hairpin structure can lead to reductions in cellular levels of the viral RNA due to reductions in RNA stability, and it was suggested that the Poly(A) hairpin may not actually play a significant role in genome packaging.¹⁰² As shown in the present studies, removal of Poly(A) from $5'-L^{344-\Delta TAR}$ does not significantly affect its structure, dimerization, or NC binding properties. Furthermore, vector RNAs in which the TAR, Poly(A), and PBS loops are deleted were packaged as well as RNAs that contained only TAR and PBS deletions, indicating that Poly(A) does not play a significant role in RNA packaging.

The PBS element contains a sequence that binds human tRNA^{Lys}, a requirement for initiation of reverse transcription.^{103–107} Numerous secondary structural models of PBS have been proposed based on mutational studies, enzymatic probing, UV crosslinking, phylogenetic studies, and chemical modification assays (reviewed in Refs. 13,39,64,71,73,74,76, and 108). Deletions within the upper loop of the PBS generally do not significantly impair genome packaging, whereas the lower stem loop is required for RNA dimerization and efficient packaging.^{77,109,110} Our PBS deletion mutant was designed to maintain native base pairing in the lower stem of the PBS loop structure. Significantly, the NMR data presented here show that the A220 C2-H signal pattern was observed in [5'-L³⁴⁴]₂ and all the PBS upper loop truncated constructs, indicating that the lower stem of the PBS element adopts the predicted secondary structure. Removal of the upper PBS loop structure eliminated ~4 high-affinity NC binding sites but did not impair RNA packaging, suggesting that these NC sites may play roles in tRNA binding and/or reverse transcription but are not critical for packaging.

In vitro biochemical studies and phylogenetic predictions led to the proposal that loop residues of PolyA base pair with residues G443–C449 in the matrix encoding region of gag.¹¹¹ These putative long-range interactions have been supported by chemical probing and modeling studies, 64,75,76 but their role in packaging has not been established. Early packaging studies suggested that efficient HIV-1 genome packaging requires residues within the first several hundred nucleotides of gag,¹¹² and could therefore be dependent on Poly(A)-gag interactions. However, the present studies show that RNAs containing the native 5'-L (nucleotides 1–355, which include residues required for U5:AUG base pairing) but missing nearly the entire gag open reading frame are packaged with the same efficiency as RNAs containing the intact gag gene. Thus, if the Poly(A)-gag interactions exist in vivo, they likely play roles in functions unrelated to RNA packaging.

The present findings also provide insights into the number and general locations of the high-affinity NC binding sites. Early studies using UV crosslinking indicated that NC is the only major protein that binds to the genomic RNA in HIV-1 virions, and it was estimated that approximately 15 NC molecules are bound per RNA strand.¹¹³ In contrast, in vitro cross-linking experiments involving a fragment of the 5'-L spanning from DIS through residue 170 of gag were interpreted to indicate that only $\sim 2-3$ NC or Gag proteins bind to each RNA strand,¹¹⁴ and RNA footprinting studies performed with a 5'-L construct that included the intact 5'-untranslated region and 66 nucleotides of gag were similarly interpreted in terms of only two high-affinity NC or Gag binding sites.⁷³ In addition, quantitative filterbinding and ITC assays indicated that NC can bind with high affinity ($K_d \approx 55-400$ nM) to fragments of the HIV-1 5'-L between the DIS and AUG hairpins.⁴³⁻

^{45,66,77} Our ITC data of 5'-L and the truncated mutants' titration with NC show that they all have high-affinity NC binding sites with K_d in the same range (Table 1). Our current and recent studies⁴⁶ are most consistent with the earlier work of Darlix *et al.*¹¹³ and indicate that the 5'-L contains approximately 16 high-affinity NC binding sites per RNA strand. The locations of the high-affinity NC binding sites of zinc ejection on SHAPE activity.⁷⁶ These studies suggested that the 5'-L contains 7 NC binding sites, none of which were proposed to reside within the TAR, Poly(A), or upper PBS loops.⁷⁶ Although these results are incompatible with our present findings, we cannot rule out the

possibility that structural changes associated with viral maturation, or constraints within the context of the packaged RNA, could lead to altered NC binding.

Sakuragi et al. have used a novel approach to identify the minimal determinants of HIV-1 genome dimerization and packaging that involves duplication of the 5'-L in downstream (ectopic) regions of the RNA, which can lead to the unusual packaging of monomeric RNAs (in addition to dimeric RNAs).¹¹⁵ In this system, RNAs with ectopic sequences lacking the Poly(A) and SD hairpins (to prevent unwanted splicing and polyadenylation) were efficiently packaged as monomers.¹¹⁵ This group subsequently showed that RNAs lacking the native 5'-L, but containing two downstream copies of a sequence that includes TAR through the first half of gag, but that lacks the Poly(A) and PBS hairpins (a 19-nucleotide deletion in the upper loop to eliminate ectopic initiation of reverse transcription), exclusively package monomeric RNAs, although virion production and packaging efficiency were reduced to $\sim 20\%$ and $\sim 63\%$ of wild-type levels, respectively.¹¹⁶ Using a similar packaging system that included a much smaller portion of gag, these investigators showed that deletions in the upper loop of the PBS do not significantly affect dimerization or packaging, but deletions in U5 and/ or the lower PBS stem lead to significant reductions in packaging.¹¹⁰ Surprisingly, severe modifications and deletions within the DIS hairpin led to relatively modest reductions in packaging levels (to 18-50% of WT levels).¹¹⁰ A systematic deletion mutagenesis study revealed that relatively efficient packaging of monomeric RNAs can be achieved by RNAs that contain a fully native 5'-L and an ectopic element that lacks the TAR, Poly(A), upper PBS loop, and SD sequences. 108 Since the RNAs employed contained a native 5'-L sequence, these studies do not rule out the possibility that packaging could depend on the presence of at least one fully native leader RNA per dimer. However, the findings are nevertheless fully compatible with our observation that, under the experimental conditions used here, relatively efficient packaging can be achieved by RNAs that lack the TAR, Poly(A), and upper PBS loops.

In summary, our findings identify 5'-L^{344-ΔTAR- Δ Poly(A)- Δ PBS as an essential HIV-1 core encapsidation element. This fragment of the 5'-L exhibits ¹H NMR spectral features similar to those of the intact 5'-L, possesses native-like *in vitro* dimerization and NC binding properties, and is required for promoting RNA packaging into VLPs. Packaging deficiencies observed here and by others^{48–50,98} upon deletion of TAR are likely due to misfolding of the monomeric form of the 5'-L. The clustering of high-affinity NC binding sites, coupled with the known self-association propensities of the CA domain of Gag, could function synergistically to promote cooperative assembly of the Gag:RNA complex that is trafficked} to the plasma membrane and nucleates virus assembly. Since the minimal packaging signal comprises only 159 nucleotides, high-resolution structural studies by NMR or combined NMR/ cryo-electron tomographic approaches¹¹⁷ should now be feasible.

Materials and Methods

Plasmid construction

The 5'-L, 5'-L^{U5:AUG}, and 5'-L^{HP-AUG} mutants were cloned into PUC19 as previously described. ⁴⁶ 5'-L³⁴⁴ and $5'-L^{344-\Delta TAR}$ were subcloned from the 5'-L plasmid using 5'-CGG TCG AAT TCT AAT ACG ACT CAC TAT AGG TCT CTC TGG TTA GAC CAC ATC TGA GCC TGG GAG CTC TCT GGC-3' and 5'-CGG TCG AAT TCT AAT ACG ACT CAC TAT AGG CAC TGC TTA AGC CTC AAT AAA GCT TGC CTT GAG ATC-3' as forward primer, respectively, and 5'-CCC TAG GAT CCC CCG GGC GCA CCC ATC TCT CTC CTT CTA-3' as their mutual reverse primer. $5'-L^{344-\Delta PBS}$ was made by MegaPrimer mutagenesis method using $5'-L^{344}$ as template. The four primers used are 5'-CGG TCG AAT TCT AAT ACG ACT CAC TAT AGG TCT CTC TGG TTA GAC CAG ATC-3', 5'-CCT GCG TCG AGA GAT CTC CTC TGG TCT CCC AGA GTC ACA CAA CAG ACG GGC ACA CAC TAC TTT GAG CA-3', 5'-GAG ACC AGA GGA GAT CTC TCG ACG CAG GAC TCG GCT TGC TGG AGA CGG CAA GAG GCG AGG GGC GGC GAC-3', and 5'-CCC TAG CTA GCT TCT GCC TCC GCT AGT CAA AAT TTT TGG CGT ACT CAC CA-3'. 5'-L^{344- Δ TAR- Δ PBS and 5'-L^{344- Δ TAR- Δ Poly(A)- Δ PBS were then cloned from 5'-L^{344- Δ PBS plasmid using the forward}}} primer 5'-CGG TCG AAT TCT AAT ACG ACT CAC TAT ÂGG CAC TGC TTA AGC CTC AAT AAA GCT TGC CTT GAG ATC-3' and 5'-CGG TCG AAT TCT AAT ACG ACT CAC TAT AGG TGT GCC CGT CGT TGT GTG ACT CTG GTG AGA GCC AGA GGA GA-3', respectively, and the mutual reverse primer 5'-CCC TAG GAT CCC CCG GGC GCA CCC ATC TCT CTC CTT CTA-3'. All of the mutations were sent to the DNA sequencing facility (University of Maryland, Baltimore County) to ensure that the desired mutations were obtained. The packaging test vector 5[']-L^{NL4-3} (Native vector; also

The packaging test vector 5'-L^{NL4-3} (Native vector; also called pSRK1428-4) was generated by incorporating 689 bases of NL4-3 5'-L region sequences in place of the corresponding sequences in HIV-lac.¹¹⁸ Variants with 5'-L^{Δ TAR</sub>, 5'-L^{Δ TAR- Δ PBS</sub>, or 5'-L^{Δ TAR- Δ Poly(A)- Δ PBS</sub> mutations were constructed by introducing mutations into the NL4-3 5'-L region using overlap extension PCR, and using these to replace the corresponding regions of 5'-L^{NL4-3}. Correct structures of all plasmids were confirmed by sequencing. For the initial series of test vectors, a stop codon was introduced after the 85th codon in the *gag* open reading frame. For *gag* deletion mutants, PCR mutagenesis was used to delete sequences downstream of the 15th nucleotide in *gag* and fuse the 15th residue to the NotI restriction site of HIV-lac.¹¹⁸ Riboprobe RNAs to detect helper and test vector RNAs were templated by pEG676-1, a derivative of pBluescript KS (Stratagene), which includes sequences complementary to 286 bases of CMV (cytomegalovirus) promoter shared by helper and test}}} vectors plus additional 14 bases sequences unique to the test vectors, preceded by a T7 promoter.

RNA packaging experiments

Human 293T cells were cultured and maintained in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% fetal bovine serum (Gemini). Cells were grown at 37 °C under 5% CO2. To generate virions for RNA packaging analysis, we co-transfected 293T cells with a total of 20 µg of plasmid DNA that included one test vector plasmid and pCMV Δ R8.2, a Ψ^- HIV-1 helper vector driven by the CMV promoter 87, at an approximate 1:3 molar ratio on 100-mm plates. Modest, empirically determined variation in co-transfection ratios was introduced for non-native vectors, to achieve similar test/ helper expression ratios for all samples. For transfection by polyethyleneimine, plasmid DNA was mixed with 4 μ g of polyethyleneimine (Polysciences) per microgram of DNA in 1 ml of 150 mM NaCl by vortexing at medium speed for 10 s. After room temperature incubation for 15 min, the mixture was added dropwise to medium on cells. The transfection mixture was replaced with fresh Dulbecco's modified Eagle's medium 24 h after transfection.

Tissue culture medium was harvested at 48 h posttransfection, pooled, and filtered through 0.2-μm-poresize filters. Virus-containing medium was concentrated by centrifugation through a 2-ml 20% sucrose cushion in phosphate-buffered saline for 2 h at 4 °C and 25,000 rpm using a Sorvall Surespin 630 rotor in a Sorvall Discovery 90 ultracentrifuge. Viral pellets were suspended in 0.5 ml of Trizol reagent (Invitrogen), and RNA was isolated according to the manufacturer's instructions. Cellular RNA was obtained by harvesting cells 48 h after transfection by scraping cells into 2 ml of Trizol reagent per 100-mm plate, and RNA was isolated according to the manufacturer's instructions.

RPAs were performed and quantified by phosphorimager as previously described.90 Riboprobes templated by pEG676-1 (see above) were used in Figs. 4 and 5 experiments. This riboprobe was designed to contain 14 bases unique to the helper as well as 268 shared bases of complementarity, to allow simultaneous detection of both vector and helper RNAs with a single probe. The experiment shown in Fig. 4b included a second riboprobe templated by pBru7SL,89 which detects 7SL RNA. The signals in the cluster of bands detected with the 7SL probe were summed: the appearance of multiple bands is typical and results from allelic heterogeneity in this multi-copy gene and some probe nicking.85 For each experiment, the amount of riboprobe loaded in the undigested probe lane was 1% of that used in the hybridization reaction for each sample lane. Protein was removed from the aqueous phase by Trizol (confirmed by spectrophotometric analysis of the RNA).

RNA in vitro transcription

The pUC19 vectors containing different RNA clones were transformed into DH5 α cells. Large-scale DNA template preparation was performed using QIAGEN Plasmid Mega Kit (Qiagen). *In vitro* transcription by T7 polymerase¹¹⁹ is performed in solution with a mixture of transcription buffer, MgCl₂, NTPs, and DNA template.

The amount of $MgCl_2$ and NTPs was optimized for each RNA sample in order to guarantee best RNA yield. The RNA was then purified using 6% denaturing PAGE gels and recovered via Elutrap Electroelution System (Whatman).

NC purification

The 55-residue NC protein (from HIV-1 strain NL4–3) was expressed from the bacteria vector pRD2, which was transformed into BL21 (DE3) pLysE. The protein was purified under non-denaturing conditions as previously described.¹²⁰

Dimerization assay

RNA was prepared in 10 mM Tris–HCl first, heat denatured for 3 min, and cooled on ice; $10 \times ITC$ salt was then added to the RNA sample to reach a final salt concentration of 140 mM KCl, 10 mM NaCl, and 1 mM MgCl₂. After incubating at 37 °C for 1, 2, 4, 24, and 48 h, the RNA samples were loaded to 2% agarose gel prestained with ethidium bromide and run at 110 V in 1 × TB buffer (44.5 mM Tris–boric acid, pH 7.5) on ice to prevent thermal denaturation during electrophoresis.

Isothermal titration calorimetry

ITC experiments were carried out using a VP-ITC MicroCalorimeter (MicroCal Corp., Northampton, MA; 1989). NC (180 $\mu\dot{M})$ in ITC buffer (10 mM Tris-HCl, pH 7.0, 140 mM NaCl, 10 mM NaCl, 1 mM MgCl₂, and 100 μM β-mercaptoethanol) was loaded into the injection syringe. The calorimetry cell was loaded with RNA (0.8 µM). The RNA sample (in 10 mM Tris-HCl) was heat denatured and then cooled on ice; 10 × ITC salt was then added to reach a final salt concentration of 140 mM KCl, 10 mM NaCl, and 1 mM MgCl₂. The RNA sample was then incubated overnight at 37 °C before loading into the calorimetry cell. After thermal equilibration at 30 °C and the initial 60-s delay, 28 serial injections of NC protein (10 µl each) were made into the calorimetry cell. The thermodynamic parameters of the NC-RNA interaction were determined by a nonlinear squares fit of the data using two-site binding mode.

Acknowledgements

This work was supported by a grant from the National Institute of General Medical Sciences (GM 42561). K.E. was supported by a grant from the National Institute of General Medical Sciences for enhancing minority access to research careers (MARC U*STAR 2T34 GM008663). We thank the Howard Hughes Medical Institute staff at the University of Maryland Baltimore County for technical assistance.

References

- Hu, W.-S. & Temin, H. M. (1990). Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl Acad. Sci. USA*, 87, 1556–1560.
- 2. Hu, W. S. & Temin, H. M. (1990). Retroviral recombination and reverse transcription. *Science*, **250**, 1227–1233.
- 3. Johnson, S. F. & Telesnitsky, A. (2010). Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLoS Pathog.* 6.
- Gratton, S., Cheynier, R., Dumaurier, M. J., Oksenhendler, E. & Wain-Hobson, S. (2000). Highly restricted spread of HIV-1 and multiply infected cells within splenic germinal centers. *Proc. Natl Acad. Sci. USA*, 97, 14566–14571.
- Jung, A., Maier, R., Vartanian, J. P., Bocharov, G., Jung, V., Fischer, U. *et al.* (2002). Recombination: multiply infected spleen cells in HIV patients. *Nature*, 418, 144.
- Quinones-Mateu, M. E. & Arts, E. J. (1999). Recombination in HIV-1: update and implications. *AIDS Rev.* 1, 89–100.
- McCutchan, F. E. (2006). Global epidemiology of HIV. J. Med. Virol. 78, S7–S12.
- Rein, A. (1994). Retroviral RNA packaging: a review. Arch. Virol. 9, 513–522.
- Berkowitz, R., Fisher, J. & Goff, S. P. (1996). RNA packaging. Curr. Top. Microbiol. Immun. 214, 177–218.
- 10. Greatorex, J. & Lever, A. (1998). Retroviral RNA dimer linkage. J. Gen. Virol. **79**, 2877–2882.
- Paillart, J.-C., Marquet, R., Skripkin, E., Ehresmann, C. & Ehresmann, B. (1996). Dimerization of retroviral genomic RNAs: structural and functional implications. *Biochimie*, 78, 639–653.
- Jewell, N. A. & Mansky, L. M. (2000). In the beginning: genome recognition, RNA encapsidation and the initiation of complex retrovirus assembly. J. Gen. Virol. 81, 1889–1899.
- Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R. & Mak, J. (2004). Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Microbiol.* 2, 461–472.
- Russell, R. S., Liang, C. & Wainberg, M. A. (2004). Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? *Retrovirology*, 1.
 Greatorex, J. (2004). The retroviral RNA dimer
- Greatorex, J. (2004). The retroviral RNA dimer linkage: different structures may reflect different roles. *Retrovirology*, 1.
- D'Souza, V. & Summers, M. F. (2005). How retroviruses select their genomes. *Nat. Rev. Microbiol.* 3, 643–655.
- Jouvenet, N., Laine, S., Pessel-Vivares, L. & Mougel, M. (2011). Cell biology of retroviral RNA packaging. *RNA Biol.* 1.
- Darlix, J. L., Godet, J., Ivanyi-Nagy, R., Fosse, P., Mauffret, O. & Mely, Y. (2011). Flexible nature and specific functions of the HIV-1 nucleocapsid protein. *J. Mol. Biol.* **410**, 565–581.
- Chin, M. P. S., Rhodes, T. D., Chen, J., Fu, W. & Hu, W.-S. (2005). Identification of a major restriction in

HIV-1 intersubtype recombination. *Proc. Natl Acad. Sci. USA*, **102**, 9002–9007.

- Moore, M. D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R. G. & Hu, W.-S. (2007). Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. J. Virol. 81, 4002–4011.
- Rhodes, T., Wargo, H. & Hu, W.-S. (2003). High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. *J. Virol.* 77, 11193–11200.
- 22. Onafuwa, A., An, W., Robson, N. D. & Telesnitsky, A. (2003). Human immunodeficiency virus type 1 genetic recombination is more frequent than that of Moloney murine leukemia virus despite similar template switching rates. J. Virol. 77, 4577–4587.
- Moore, M. D., Nikolaitchik, O. A., Chen, J., Hammarskjold, M. L., Rekosh, D. & Hu, W. S. (2009). Probing the HIV-1 genomic RNA trafficking pathway and dimerization by genetic recombination and single virion analyses. *PLoS Pathog.* 5, e1000627.
- Jouvenet, N., Bieniasz, P. D. & Simon, S. M. (2008). Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature*, 454, 236–240.
- Jouvenet, N., Simon, S. M. & Bieniasz, P. D. (2009). Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc. Natl Acad. Sci. USA*, **106**, 19114–19119.
- Bender, W., Chien, Y.-H., Chattopadhyay, S., Vogt, P. K., Gardner, M. B. & Davidson, N. (1978). High-molecular-weight RNAs of AKR, NZB and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. *J. Virol.* 25, 888–896.
- Kung, H.-J., Hu, S., Bender, W., Baily, J. M., Davidson, N., Nicolson, M. O. & McAllister, R. M. (1976). RD-114, baboon and wolly monkey viral RNAs compared in size and structure. *Cell*, 7, 609–620.
- Maisel, J., Bender, W., Hu, S., Duesberg, P. H. & Davidson, N. (1978). Structure of 50 to 70S RNA from Moloney sarcoma viruses. J. Virol. 25, 384–394.
- Murti, K. G., Bondurant, M. & Tereba, A. (1981). Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. *J. Virol.* 37, 411–419.
- Fu, W. & Rein, A. (1993). Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J. Virol.* 67, 5443–5449.
- Oritz-Conde, B. A. & Hughes, S. H. (1999). Studies of the genomic RNA of leukosis viruses: implications for RNA dimerization. *J. Virol.* **73**, 7165–7174.
 Rice, W. G., Schaeffer, C. A., Harten, B., Villinger, F.,
- Rice, W. G., Schaeffer, C. A., Harten, B., Villinger, F., South, T. L., Summers, M. F. *et al.* (1993). Inhibition of HIV-1 infectivity by zinc-ejecting aromatic C-nitroso compounds. *Nature*, **361**, 473–475.
- Rice, W. G., Turpin, J. A., Schaeffer, C. A., Graham, L., Clanton, D., Buckheit, R. W. *et al.* (1996). Evaluation of selected chemotypes in coupled cellular and molecular target-based screens identifies novel HIV-1 zinc finger inhibitors. *J. Med. Chem.* 39, 3606–3616.

- Grigorov, B., Bocquin, A., Gabus, C., Avilov, S., Mely, Y., Agopian, A. *et al.* (2011). Identification of a methylated oligoribonucleotide as a potent inhibitor of HIV-1 reverse transcription complex. *Nucleic Acids Res.* 39, 5586–5596.
- Avilov, S. V., Boudier, C., Gottikh, M., Darlix, J. L. & Mely, Y. (2011). Characterization of the inhibition mechanism of the HIV-1 nucleocapsid protein chaperone activities by methylated oligoribonucleotides. *Antimicrob. Agents Chemother.* 56, 1010–1018.
- Goldschmidt, V., Miller, L. M., Dde Rocquigny, H., Darlix, J.-L. & Mely, Y. (2010). The nucleocapsid protein of HIV-1 as a promising therapeutic target for antiviral drugs. *HIV Ther.* 4, 179–198.
- De Guzman, R. N., Turner, R. B. & Summers, M. F. (1999). Protein–RNA recognition. *Biopolymers*, 48, 181–195.
- 38. Lever, A. M. (2007). HIV-1 RNA packaging. *Adv. Pharmacol.* 55, 1–32.
- Lu, K., Heng, X. & Summers, M. F. (2011). Structural determinants and mechanism of HIV-1 genome packaging. J. Mol. Biol. 410, 609–633.
- D'Souza, V., Dey, A., Habib, D. & Summers, M. F. (2004). NMR structure of the 101 nucleotide core encapsidation signal of the Moloney murine leukemia virus. J. Mol. Biol. 337, 427–442.
- D'Souza, V. & Summers, M. F. (2004). Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature*, 431, 586–590.
- Zhou, J., Bean, R. L., Vogt, V. M. & Summers, M. F. (2006). Solution structure of the Rous sarcoma virus nucleocapsid protein:uY RNA packaging signal complex. J. Mol. Biol. 365, 453–467.
- De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N. & Summers, M. F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 Ψ-RNA recognition element. *Science*, 279, 384–388.
- 44. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K., Wu, Z.-R. & Summers, M. F. (2000). NMR structure of the HIV-1 nucleocapsid protein bound to stem–loop SL2 of the Ψ-RNA packaging signal. J. Mol. Biol. 301, 491–511.
- Spriggs, S., Garyu, L., Connor, R. & Summers, M. F. (2008). Potential intra- and intermolecular interactions involving the Unique-5' Region of the HIV-1 5'-UTR. *Biochemistry*, 46, 13064–13073.
- Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E., Kharytonchyk, S. *et al.* (2011). NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science*, 344, 242–245.
- Nikolaitchik, O., Rhodes, T. D., Ott, D. & Hu, W.-S. (2006). Effects of mutations in the human immunodeficiency virus type 1 gag gene on RNA packaging and recombination. 80, 4691–4697.
- Helga-Maria, C., Hammarskjold, M.-L. & Rekosh, D. (1999). An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type-1 genomic RNA. J. Virol. 73, 4127–4135.
- McBride, M. S., Schwartz, M. D. & Panganiban, A. T. (1997). Efficient encapsidation of human immunodeficiency virus type 1 vectors and

further characterization of cis elements required for encapsidation. J. Virol. 71, 4544–4554.

- Das, A. T., Klaver, B. & Berkhout, B. (1998). The 5' and 3' TAR elements of human immunodeficiency virus exert effects at several points in the virus life cycle. J. Virol. 72, 9217–9223.
- Das, A. T., Harwig, A., Vrolijk, M. M. & Berkhout, B. (2007). The TAR hairpin of human immunodeficiency virus Type 1 can be deleted when not required for Tat-mediated activation of transcription. *J. Virol.* 81, 7742–7748.
- Brandt, S., Blissenbach, M., Grewe, B., Konietzny, R., Grunwald, T. & Uberla, K. (2007). Rev proteins of human and simian immunodeficiency virus enhance RNA encapsidation. *PLoS Pathog.* 3, e54.
- Poon, D. T. K., Wu, J. & Aldovini, A. (1996). Charged amino acid residues of human immunodeficiency virus type-1 nucleocapsid P7 protein involved in RNA pacakaging and infectivity. *J. Virol.* 70, 6607–6617.
- Cimarelli, A., Sandin, S., Hoglund, S. & Luban, J. (2000). Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. J. Virol. 74, 3046–3057.
- Wang, S.-W. & Aldovini, A. (2002). RNA incorporation is critical for retroviral particle integrity after cell membrane assembly of Gag complexes. *J. Virol.* 76, 11853–11865.
- Muriaux, D., Mirro, J., Harvin, D. & Rein, A. (2001). RNA is a structural element in retrovirus particles. *Proc. Natl Acad. Sci. USA*, 98, 5246–5251.
- 57. Muriaux, D., Mirro, J., Nagashima, K., Harvin, D. & Rein, A. (2002). Murine leukemia virus nucleocapsid mutant particles lacking viral RNA encapsidate ribosomes. *J. Virol.* **76**, 11405–11413.
- Harrison, G. P. & Lever, A. M. L. (1992). The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. J. Virol. 66, 4144–4153.
- Hayashi, T., Shioda, T., Iwakura, Y. & Shibuta, H. (1992). RNA packaging signal of human immunodeficiency virus type 1. *Virology*, **188**, 590–599.
- Baudin, F., Marquet, R., Isel, C., Darlix, J.-L., Ehresmann, B. & Ehresmann, C. (1993). Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J. Mol. Biol.* 229, 382–397.
- Hayashi, T., Ueno, Y. & Okamoto, T. (1993). Elucidation of a conserved RNA stem–loop structure in the packaging signal of human immunodeficiency virus type 1. *FEBS Lett.* **327**, 213–218.
- McBride, M. S. & Panganiban, A. T. (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J. Virol.* **70**, 2963–2973.
- 63. McBride, M. S. & Panganiban, A. T. (1997). Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J. Virol.* **71**, 2050–2058.
- Damgaard, C. K., Andersen, E. S., Knudsen, B., Gorodkin, J. & Kjems, J. (2004). RNA interactions in the 5' region of the HIV-1 genome. *J. Mol. Biol.* 336, 369–379.

- Harrison, G. P., Miele, G., Hunter, E. & Lever, A. M. L. (1998). Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. *J. Virol.* 72, 5886–5896.
- Clever, J., Sassetti, C. & 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.
- 67. Clever, J. L. & Parslow, T. G. (1997). Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J. Virol.* **71**, 3407–3414.
- Berkhout, B. (1992). Structural features in TAR RNA of human and simian immunodeficiency viruses: a phylogenetic analysis. *Nucleic Acids Res.* 20, 27–31.
- Berkhout, B., Klaver, B. & Das, A. T. (1995). A conserved hairpin structure predicted for the poly(A) signal of human and simian immunodeficiency viruses. *Virology*, 207, 276–281.
- Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B. & Ehresmann, C. (1994). Identification of the primary site of the human immunodefieiency virus type 1 RNA dimerization in vitro. *Proc. Natl Acad. Sci.* USA, **91**, 4945–4949.
- Abbink, T. E. M. & Berkhout, B. (2003). A novel long distance base-pairing interaction in Human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J. Biol. Chem.* 278, 11601–11611.
- 72. Berkhout, B. (1996). Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **54**, 1–34.
- Damgaard, C. K., Dyhr-Mikkelsen, H. & Kjems, J. (1998). Mapping the RNA binding sites for human immunodeficiency virus type-1 Gag and NC proteins within the complete HIV-1 and HIV-2 untranslated leader regions. *Nucleic Acids Res.* 26, 3667–3676.
- Paillart, J. C., Dettenhofer, M., Yu, X.-F., Ehresmann, C., Ehresmann, B. & Marquet, R. (2004). First snapshots of the HIV-1 RNA structure in infected cells and in virions. *J. Biol. Chem.* 279, 48397–48403.
- Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Jr., Swanstrom, R. *et al.* (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, 460, 711–716.
- Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H. *et al.* (2008). Highthroughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* 6, 883–899.
- Clever, J. L., Miranda, J. D. & Parslow, T. G. (2002). RNA structure and packaging signals in the 5' leader region of the human immunodeficiency virus type 1 genome. *J. Virol.* **76**, 12381–12387.
- Woese, C. R., Winder, S. & Gutell, R. R. (1990). Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc. Natl Acad. Sci. USA*, 87, 8467–8471.
- Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M. & Pardi, A. (1996). A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* 264, 968–980.

- Szewczak, A. A. & Moore, P. B. (1995). The sarcin/ricin loop, a modular RNA. J. Mol. Biol. 247, 81–98.
- Szewczak, A. A., Moore, P. B., Chan, Y. L. & Wool, I. G. (1993). The conformation of the sarcin/ricin loop from 28S ribosomal RNA. *Proc. Natl Acad. Sci.* USA, 90, 9581–9585.
- Orita, M., Nishikawa, F., Kohno, T., Senda, T., Mitsui, Y., Endo, Y. *et al.* (1996). Highresolution NMR study of GdAGA tetranucleotide loop that is an improved substrate for ricin, a cytotoxic plant protein. *Nucleic Acids Res.* 24, 611–618.
- Orita, M., Nishikawa, F., Shimayama, T., Taira, K., Endo, Y. & Nishikawa, S. (1993). High-resolution NMR study of a synthetic oligonucleotide with a tetranucleotide GAGA loop that is a substrate for the cytotoxic protein, ricin. *Nucleic Acids Res.* 21, 5670–5678.
- Correll, C. C., Munishkin, A., Chan, Y., Ren, Z., Wool, I. G. & Steitz, T. A. (1998). Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc. Natl Acad. Sci. USA*, 95, 13436–13441.
- Amarasinghe, G. K., Zhou, J., Miskimon, M., Chancellor, K. J., McDonald, J. A., Matthews, A. G. *et al.* (2001). Stem–loop SL4 of the HIV-1 Ψ-RNA packaging signal exhibits weak affinity for the nucleocapsid protein. Structural studies and implications for genome recognition. *J. Mol. Biol.* 314, 961–969.
- Kanevsky, I., Chaminade, F., Ficheux, D., Moumen, A., Gorelick, R., Negroni, M. *et al.* (2005). Specific interactions between HIV-1 nucleocapsid protein and the TAR element. *J. Mol. Biol.* 348, 1059–1077.
- Rulli, S. J., Jr, Hibbert, C. S., Mirro, J., Pederson, T., Biswal, S. & Rein, A. (2007). Selective and nonselective packaging of cellular RNAs in retrovirus particles. J. Virol. 81, 6623–6631.
- Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H. *et al.* (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, 272, 263–267.
- Onafuwa-Nuga, A. A., Telesnitsky, A. & King, S. R. (2006). 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles. *RNA*, **12**, 542–546.
- Onafuwa-Nuga, A. A., King, S. R. & Telesnitsky, A. (2005). Nonrandom packaging of host RNAs in Moloney murine leukemia virus. *J. Virol.* 79, 13528–13537.
- 91. Anderson, E. S., Contera, S. A., Knudsen, B., Damgaard, C., Besenbacher, F. & Kjems, J. (2004). Role of the trans-activation response element in dimerization of HIV-1 RNA. *J. Biol. Chem.* **279**, 22243–22249.
- Song, R., Kafaie, J. & Laughrea, M. (2008). Role of the 5' TAR stem–loop and the U5-AUG duplex in dimerization of HIV-1 genomic RNA. *Biochemistry*, 47, 3283–3293.
- Huthoff, H. & Berkhout, B. (2001). Two alternating structures of the HIV-1 leader RNA. RNA, 7, 143–157.

- 94. Huthoff, H. & Berkhout, B. (2001). Mutations in the TAR hairpin affect the equilibrium between alternative conformations of the HIV-1 leader RNA. *Nucleic Acids Res.* **29**, 2594–2600.
- 95. Song, R., Kafaie, J., Yang, L. & 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, 1084–1096.
- Berkhout, B., Silverman, R. H. & Jeang, K. T. (1989). Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell*, 59, 273–282.
- Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J. *et al.* (1989). Human immunodeficiency virus 1 tat protein binds trans-activationresponsive region (TAR) RNA in vitro. *Proc. Natl Acad. Sci. USA*, 86, 6925–6929.
- Clever, J. L., Eckstein, D. A. & Parslow, T. G. (1999). Genetic dissociation of the encapsidation and reverse transcription functions in the 5' R region of human immunodeficiency virus type 1. J. Virol. 73, 101–109.
- Höglund, S., Öhagen, A., Goncalves, J., Panganiban, A. T. & Gabuzda, D. (1997). Ultrastructure of HIV-1 genomic RNA. *Virology*, 233, 271–279.
- 100. Russell, R. S., Hu, J., Laughrea, M., Wainberg, M. A. & Liang, C. (2002). Deficient dimerization of human immunodeficiency virus type 1 RNA caused by mutations of the U5 RNA sequences. *Virology*, **303**, 152–163.
- 101. Das, A. T., Klaver, B., Klasens, B. I., van Wamel, J. L. & Berkhout, B. (1997). A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. *J. Virol.* **71**, 2346–2356.
- 102. Das, A. T., Klaver, B. & Berkhout, B. (1999). A hairpin structure in the R region of the human immunodeficiency virus type 1 RNA genome is instrumental in polyadenylation site selection. *J. Virol.* **73**, 81–91.
- Ísel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B. & Marquet, R. (1996). Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA3Lys. *EMBO J.* 15, 917–924.
 Huang, Y., Khorchid, A., Wang, J., Parniak, M. A.,
- 104. Huang, Y., Khorchid, A., Wang, J., Parniak, M. A., Darlix, J. L., Wainberg, M. A. & Kleiman, L. (1997). Effect of mutations in the nucleocapsid protein (NCp7) upon Pr160(gag-pol) and tRNA(Lys) incorporation into human immunodeficiency virus type 1. *J. Virol.* **71**, 4378–4384.
- 105. Zhang, Z., Yu, Q., Kang, S. M., Buescher, J. & Morrow, C. D. (1998). Preferential completion of human immunodeficiency virus type 1 proviruses initiated with tRNA3Lys rather than tRNA1,2Lys. *J. Virol.* **72**, 5464–5471.
- 106. Beerens, N. & Berkhout, B. (2002). The tRNA primer activation signal in the human immunode-ficiency virus type 1 genome is important for initiation and processive elongation of reverse transcription. *J. Virol.* **76**, 2329–2339.
- 107. Goldschmidt, V., Rigourd, M., Ehresmann, C., Le Grice, S. F., Ehresmann, B. & Marquet, R. (2002). Direct and indirect contributions of RNA secondary

structure elements to the initiation of HIV-1 reverse transcription. *J. Biol. Chem.* **277**, 43233–43242.

- Sakuragi, J., Sakuragi, S. & Shioda, T. (2007). Minimal region sufficient for genome dimerization in the human immunodeficiency virus type 1 virion and its potential roles in the early stages of viral replication. *J. Virol.* 81, 7985–7992.
- 109. Vicenzi, E., Dimitrov, D. S., Engelman, A., Migone, T. S., Purcell, D. F., Leonard, J. *et al.* (1994). An integration-defective U5 deletion mutant of human immunodeficiency virus type 1 reverts by eliminating additional long terminal repeat sequences. *J. Virol.* **68**, 7879–7890.
- Sakuragi, J.-I., Ueda, S., Iwamoto, A. & Shioda, T. (2003). Possible role of dimerization in human immunodeficiency virus type-1 genome RNA packaging. *J. Virol.* 77, 4060–4069.
- 111. Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. (2002). In vitro evidence for a long range pseudoknot in the 5'-untranslated and matrix coding regions of the HIV-1 genomic RNA. J. Biol. Chem. 277, 5995–6004.
- 112. Berkowitz, R. D., Hammarskjold, M.-L., Helga-Maria, C., Rekosh, D. & Goff, S. P. (1995). 5' Regions of HIV-1 RNAs are not sufficient for encapsidation: implications for the HIV-1 packaging signal. *Virology*, 212, 718–723.
- 113. Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F. & Barre-Sinoussi, F. (1990). Cis elements and transacting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* **216**, 689–699.

- 114. Roldan, A., Russell, R. S., Marchand, B., Gotte, M., Liang, C. & Wainberg, M. A. (2004). In vitro identification and characterization of an early complex linking HIV-1 genomic RNA recognition and Pr55Gag multimerization. *J. Biol. Chem.* 279, 39886–39894.
- 115. Sakuragi, J.-I., Shioda, T. & Panganiban, A. T. (2001). Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. J. Virol. 75, 2557–2565.
- Sakuragi, J.-I., Iwamoto, A. & Shioda, T. (2002). Dissociation of genome dimerization from packaging functions and virion maturation of human immunodeficiency virus type 1. *J. Virol.* 76, 959–967.
- 117. Miyazaki, Y., Irobalieva, R. N., Tolbert, B. S., Smalls-Manty, A., Iyalla, K., Loeliger, K. *et al.* (2010). Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography. *J. Mol. Biol.* **404**, 751–772.
- An, W. & Telesnitsky, A. (2001). Frequency of direct repeat deletion in a human immunodeficiency virus type 1 vector during reverse transcription in human cells. *Virology*, 286, 475–482.
- Milligan, J. F. & Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* 180, 51–62.
- Lee, B. M., De Guzman, R. N., Turner, B. G., Tjandra, N. & Summers, M. F. (1998). Dynamical behavior of the HIV-1 nucleocapsid protein. *J. Mol. Biol.* 279, 633–649.