Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 3393–3397, April 1993
Biochemistry

Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA

WESLEY I. SUNDQUIST* AND SHAUN HEAPHY†

*Department of Biochemistry, University of Utah, Salt Lake City, UT 84132; and †Department of Microbiology, University of Leicester, Leicester LE1 9HN, England

Communicated by Aaron Klug, December 21, 1992 (received for review September 21, 1992)

ABSTRACT **Retroviruses package two homologous single-stranded RNA genomes within a gag protein–RNA complex. In mature virion particles, the two RNA strands are thought to associate primarily through direct RNA–RNA interactions, although the structural basis for this stable association is unknown. We now report that a 127-nucleotide (nt) HIV-1_{NL4-3} RNA fragment (positions 732–858) encompassing the 5′ end of the *gag* gene dimerizes spontaneously under high ionic strength conditions in the absence of any protein cofactor. The HIV-1 RNA dimer is dramatically and specifically stabilized by the monovalent cation potassium. Thermal dissociation of the dimer occurs at 80°C in 100 mM K^+ (5 mM Mg^{2+}) but at significantly lower temperatures in the presence of either smaller or larger monovalent cations (100 mM Li^+, 40°C; 100 mM Na^+, 55°C; 100 mM Cs^+, 30°C). Deletion analyses of the 3′ end of the 127-nt fragment reveal that an HIV-1 RNA fragment as short as 94 nt (732–825) can dimerize spontaneously, but a further 9-base deletion of the purine-rich sequence, GGGGGAGAA from positions 817 through 825, eliminates dimerization. These experimental results support a model in which HIV-1 RNA dimerizes by forming an interstrand quadruple helix stabilized by guanine (and/or purine)-base tetrads in analogy to the well-known dimerization of telomeric DNA. We speculate that this structure may also mediate the association of genomic HIV-1 RNA *in vivo*, revealing how RNA itself can achieve the self-recognition required for subsequent genetic recombination.**

The organization of genomic retroviral RNA into infectious virion particles constitutes an important but poorly understood step in the infectious cycle of retroviruses (reviewed in ref. 1). Mature retroviral virions contain two homologous copies of their single-stranded RNA genomes, which are stably associated within an RNA–gag protein complex. This intimate association of two homologous genomic strands accounts in part for the remarkable genetic variability of retroviruses because information on both strands can be utilized upon subsequent infection (2–4). The most stable contact between the two genomic RNAs, termed the "dimer linkage structure" (DLS), occurs at a discrete site near the 5′ end of each strand (5). The two strands emanate from the DLS with the same orientation (i.e., 5′ to 5′ and 3′ to 3′), and this parallel alignment may facilitate subsequent recombination via a reverse transcriptase strand switching mechanism (2, 3).

There is strong genetic and biochemical evidence that genomic RNA dimerization in a number of different retroviruses is facilitated by interactions with viral gag protein(s) (6–18). Nevertheless, the two RNA strands in the dimer linkage structure appear to associate primarily via direct RNA–RNA interactions, since genomic RNA dimers extracted from virions are quite stable under nondenaturing conditions even after all proteins are removed (5, 19, 20). Furthermore, several retroviral RNA genomes, including human immunodeficiency virus 1 (HIV-1), have been shown to dimerize spontaneously in the absence of protein cofactors under conditions of high ionic strength (7, 9, 21, 22). HIV-1_{Mal} sequences sufficient to support spontaneous dimerization have been mapped to a 113-nucleotide (nt) domain (corresponding to positions 743–855 in HIV-1_{NL4-3}) encompassing the 5′ end of the *gag* gene (21). The structural motif mediating the association of two identical (rather than complementary) single strands within the DLS has remained unknown, however.

Another well-characterized example of identical nucleic acid strands forming stable molecular dimers is the self-association of telomeric DNA sequences (reviewed in refs. 23–25). We (26) and others (27–30) have shown that dimerization of telomeric DNA occurs via the formation of unusual intermolecular quadruple helical structures that are stabilized by guanine base tetrads (G-tetrads; Fig. 1 *Upper*). Dimerization is favored by sequences that contain two separate tracts of guanines that can fold back on themselves and dimerize into an antiparallel quadruple helix. The most remarkable feature of these guanine hairpin dimer structures is their dramatically enhanced stability in the presence of potassium, which arises as a result of size-specific chelation of K^+ within the axial channel of the quadruple helix (Fig. 1 *Lower*) (26, 31, 33–35). Molecular details of the G-tetrad hydrogen-bonding scheme, nucleotide chain configurations, and metal-binding sites have recently been revealed by high-resolution NMR and x-ray crystallographic studies (31, 36). Although single-stranded overhanging sequences capable of forming related structures are broadly conserved across eukaryotic telomeres, their biological function remains uncertain (24, 37, 38).

Self-association of retroviral genomes requires the analogous annealing of two identical nucleic acid strands, raising the possibility that retroviral RNA dimerization may also be mediated by G-tetrad formation. Furthermore, retroviral dimerization domains are purine rich, and the kinetics and ionic strength dependencies of dimerization are consistent with quadruplex formation (22). We now present biochemical evidence indicating that interstrand G-tetrads mediate the dimerization of HIV-1 RNA *in vitro*. We suggest that an analogous structure may mediate RNA dimerization *in vivo*, which would provide the first well-characterized biological function for a quadruple helical nucleic acid structure.

MATERIALS AND METHODS

RNA Synthesis and Purification. RNA molecules used in dimerization experiments were transcribed *in vitro* from linearized plasmid templates (39). Initially, HIV-1 sequences were amplified from a genomic clone of NL4-3 HIV-1 DNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DLS, dimer linkage structure; HIV-1, human immunodeficiency virus 1; G-tetrad, guanine base tetrad; nt, nucleotide.

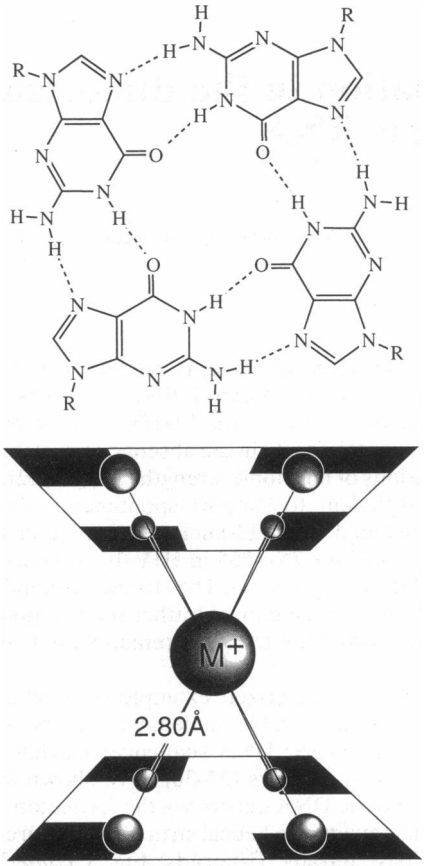


FIG. 1. (*Upper*) Hydrogen bonding scheme of the G-tetrad. The polynucleotide strands that compose the quadruple helix can adopt either a parallel or antiparallel orientation. An isostructural adenine tetrad stabilized by only four hydrogen bonds (N6 to N1) is also possible. (*Lower*) Schematic illustration of the potassium binding site within a guanine quadruple helix. The eight-coordinate chelation cage created within the axial channel of the quadruple helix by stacking of adjacent G-tetrads is shown schematically. Each guanine (shaded) surrounding the cage contributes a single O6 ligand (spheres). For clarity, the stacked tetrads are depicted without helical twist, although known structures exhibit right-handed helical twist with a helical repeat of 11.5 tetrads per turn (31, 32). The average M—O distance (2.80 Å) within the cage agrees well with the expected K—O bond length (2.78 Å) as calculated from the sum of the appropriate van der Waals radii. The M—O distance of 2.80 Å is derived from a fiber diffraction model for the quadruple helical structure of poly(G), although individual cages are likely to be somewhat heterogeneous in size (31, 32).

(40) by using PCR primers designed to introduce terminal *Eco*RI and *Hin*dIII cloning sites and an internal phage T7 RNA polymerase promoter sequence immediately upstream of the desired template sequence. PCR products were cloned into the *Eco*RI/*Hin*dIII site of pUC119, and the resulting plasmids were confirmed by DNA sequencing (41).

Transcription reaction mixtures (500 μl) containing 40 mM Tris·HCl (pH 7.9 at 25°C), 8 mM $MgCl_2$, 1 mM dithiothreitol, 2 mM spermidine, 800 μM each NTP, 50 μg of bovine serum albumin per ml, 25 μg of *Hin*dIII-linearized plasmid, 400 units of RNasin (Promega), and 600 units of T7 RNA polymerase (Pharmacia) were incubated for 2 hr at 37°C. RNA molecules were radioactively labeled either by including [α-^{32}P]UTP in the transcription mixture [2–4 μl, 10 μCi (370 kBq)/ml; Fig. 2] or by end-labeling with [γ-^{32}P]ATP (Fig. 3) and polynucleotide kinase (after removal of the 5′ triphosphate with calf intestinal phosphatase) (41). RNA transcripts were purified by PAGE (7 M urea), eluted into SDS/TE buffer (10 mM Tris/1 mM EDTA/0.5% SDS, pH 8.0), and quantitated by scintillation counting and absorption spectroscopy (42). All sequences have been converted to the HIV-1_{NL4-3} numbering scheme given in ref. 43.

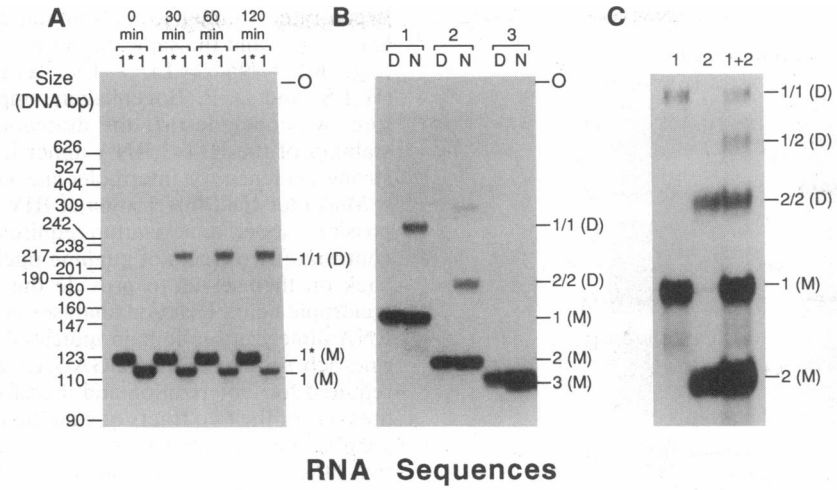
***In Vitro* Dimerization.** To promote dimerization, RNA molecules were incubated at a concentration of 1.5 μM RNA strand for 60 min at 37°C in a buffer containing 10 mM Tris (pH 7.9), 250 mM NaCl, 50 mM $MgCl_2$, and 0.1% SDS. Prior to electrophoresis, samples were diluted 1:10 into a buffer containing 5% (vol/vol) glycerol, 10 mM Tris, (pH 7.9), and 25 mM KCl. Alternatively, samples to be used as denatured controls were diluted 1:10 into 70% formamide and incubated at 90°C for 1 min. Samples were electrophoresed (15 V/cm) through 5% acrylamide gels (19:1 acrylamide/*N,N*′-methylenebisacrylamide). Both gels and running buffers contained 25 mM KCl (to enhance complex stability), 90 mM Tris borate (pH 8.0), and 1 mM EDTA.

Cation-Dependent Thermal Dissociation Experiments. The cation-dependent thermal stability of the RNA 1 dimer was measured in parallel in buffers containing one of four different monovalent cations (M^+), each at 100 mM. To form the dimer, RNA 1 was incubated for 60 min at 37°C under dimerization conditions. The resulting RNA was precipitated from ethanol, washed thoroughly with 70% ethanol, redissolved in water (4°C), and redistributed into buffers containing 100 mM M^+ (Li^+, Na^+, K^+, or Cs^+), 5 mM Mg^{2+}, and 10 mM Tris·HCl (pH 8.3) to give an ≈4 μM final strand concentration. To assay thermal dissociation, RNA samples were incubated in a water bath for 5 min at each indicated temperature, and an aliquot of each sample was removed, diluted 1:10 into loading buffer, and flash-frozen. Bulk RNA samples were then incubated at the next higher temperature, and the procedure was repeated. At the end of the experiment, the aliquots were rapidly thawed, and the concentrations of intact dimer at each temperature were assayed by nondenaturing PAGE. As a control, the dimeric complex (IIA) of the telomeric DNA construct—d(GTCGACCCGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGG)·d(CCCCAACCCCAACCCCAACCCGGGTCGAC)—was also tested for cation-dependent stability in this assay (26).

RESULTS AND DISCUSSION

Spontaneous Dimerization of the HIV-1 Sense RNA Strand. Previous deletion analyses identified a 113-nt sequence (positions 743–855) spanning the 5′ end of the HIV-1_{Mal} *gag* gene, which is capable of spontaneous dimerization *in vitro* (21). A 127-nt HIV-1_{NL4-3} RNA fragment (positions 732–858, designated RNA 1; see Fig. 2 *Lower*) containing this sequence motif slowly self-associated at neutral pH and high ionic strength (250 mM NaCl/50 mM $MgCl_2$/10 mM Tris, pH 7.4) to form a dimeric complex of reduced electrophoretic mobility (Fig. 2 *Upper A*). In contrast, an antisense RNA incubated under the same conditions remained monomeric. RNA 1 dimerization is concentration dependent and proceeds with an apparent second-order rate constant k of ≈50 liter/mol·sec (although a full kinetic analysis remains to be performed). This experiment establishes that the non-Watson-Crick structure mediating dimerization requires sequences within the fragment "732–856" in the naturally occurring sense HIV-1 RNA strand, in good agreement with a previous study (21).

To delineate the 3′ boundary of the dimerization domain, a series of RNA constructs with successively longer 3′ deletions were tested for dimerization (Fig. 2 *Upper B* and data not shown). The shortest HIV-1 RNA sequence found capable of dimerization was a 94-base fragment (RNA 2, positions 732–825) and the longest RNA sequence incapable of dimerization was an 85-base fragment (RNA 3, positions 732–816) (Fig. 2 *Upper B*). This experiment revealed that sequences necessary and sufficient for dimerization are con-



RNA Sequences

RNA 1* (Antisense Control; 130 nt):

5' GGCCUUAACCGAAUUUUUUCCCAUUUAUCUAAUUCUCCCCCGCUUAAUACCGACGCUCUCGCACCCAUCUCUCUCCUUCUAGCCUCCGCUAGUCAAAAUUUUUGGCGUACUGACCAGUCGCCGCCAAGCU 3'

RNA 1 (732-858; 131 nt):

5' GGGCGGCGACUGGUGAGUACGCCAAAAAUUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCGGUAUUAAGCGGGGGAGAAUUAGAUAAAUGGGAAAAAAUUCGGUUAAGGCCAAGCU 3'

RNA 2 (732-825; 98 nt):

5' GGGCGGCGACUGGUGAGUACGCCAAAAAUUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCGGUAUUAAGCGGGGGAGAAAGCU 3'

RNA 3 (732-816; 86 nt):

5' GGGCGGCGACUGGUGAGUACGCCAAAAAUUUUGACUAGCGGAGGCUAGAAGGAGAGAGAUGGGUGCGAGAGCGUCGGUAUUAAGCU 3'

FIG. 2. (*Upper A*) Dimerization of HIV-1 sense RNA. Monomeric sense (RNA 1) or antisense (RNA 1*) strands were incubated under dimerization conditions, and aliquots were removed from the incubation mixtures after 0, 30, 60, and 120 min. Multimeric complexes were separated from the monomers (M) by nondenaturing PAGE. An intermolecular dimer is formed by the RNA 1 sense strand [denoted 1/1 (D)] but not the RNA 1* antisense strand [denoted 1* (M)]. (*Upper B*) 3′ Deletion analysis of the sequences required for RNA 1 dimerization. 3′-Terminal HIV-1 sequences were deleted from RNA 1 (control, 127 native nt + 4 nt) to create RNA 2 (94 native nt + 4 nt), and RNA 3 (85 native nt + 1 nt). The RNAs were incubated under dimerization conditions and electrophoresed under nondenaturing conditions. Samples in lanes labeled D at the top were subjected to denaturation prior to electrophoresis as described in text, whereas samples in lanes labeled N were maintained under nondenaturing conditions. The experiment reveals that RNAs 1 and 2 form dimeric complexes [denoted 1/1, (D) and 2/2 (D), respectively], whereas RNA 3 does not dimerize [denoted 3 (M), for monomer]. As shown in lane 2N, minor complexes of reduced mobility were frequently observed in the dimerization of RNA 2, which may reflect a reduction in the specificity of dimerization resulting from deletion of nucleotides 826–858. (*Upper C*) Heterodimerization of RNAs 1 and 2. Coincubation of RNAs 1 and 2 (lane 1 + 2) under dimerization conditions yields three intermolecular complexes: the homodimers RNA 1–RNA 1 (lane 1) and RNA 2–RNA 2 (lane 2) as well as the heterodimer RNA 1–RNA 2. Formation of a single heterodimeric complex [RNA 1–RNA 2, denoted 1/2 (D)] shows that the complexes formed by RNA 1 and RNA 2 are in fact dimers. (*Lower*) RNAs 1–3 with underlined nucleotides denoting nonnative sequences introduced by the *Hin*dIII restriction site.

tained within the fragment "732–825" and implicated the purine-rich, 3′ proximal sequence GGGGGAGAA from position 817 through 825 in the DLS.

The multimerization state of intermolecular complexes was established by using two RNA molecules of different length. Coincubation of RNAs 1 and 2 under conditions of high ionic strength produced the homodimeric complexes (RNA 1–RNA 1 and RNA 2–RNA 2; Fig. 2 *Upper C*) as well as a single heterodimeric complex of intermediate molecular weight (RNA 1–RNA 2). Formation of a single heterodimeric complex ruled out oligomeric states greater than two and also ruled out the possibility that lower mobility forms of RNA 1 and RNA 2 are alternative conformations of the monomeric sense RNAs.

Cation-Dependent Stability of the HIV-1 RNA Dimer. Guanine quadruple-helical structures selectively chelate K^+ within the axial channel of the helix in sites created by the eight O6 atoms of two stacked tetrads (see Fig. 1 *Lower*) (31, 33). Therefore, K^+ imparts significantly greater thermal stability to these structures than do other larger or smaller monovalent cations (26, 33–35). This effect is seen in Fig. 3 *Right* for a reference telomeric DNA dimer (26). In this case, the quadruple helical fold-back structures that serve to anneal two telomeric DNA strands have been characterized biochemically (26–30), spectroscopically (36), and crystallographically (31). As expected, thermal dissociation of the telomeric DNA dimer is strongly cation dependent, with dissociation occurring at ≈40°C in 100 mM LiCl, 55°C in 100 mM NaCl, >70°C in 100 mM KCl, and ≈40°C in 100 mM CsCl (note that temperatures of >70°C cause melting of the double-helical portion of the telomeric DNA dimer).

We anticipated that if intermolecular G-tetrads also serve to stabilize the RNA 1 homodimer, then this complex would show similar K^+-dependent thermal stability. As shown in Fig. 3 *Left*, this is indeed the case. The thermal stability of the RNA 1–RNA 1 homodimer is maximal in 100 mM K^+ ($T_d \approx$ 80°C), whereas the dimer is significantly less stable in the presence of both smaller (Li^+, $T_d \approx$ 40°C; Na^+, $T_d \approx$ 55°C) and larger (Cs^+, T_d = 30–40°C) monovalent cations. The thermal dissociation of RNA 1–RNA 1 shown in Fig. 3 was performed in the presence of 5 mM Mg^{2+} (to mimic intracellular conditions), although we observe even greater K^+-specific thermal stability in the absence of Mg^{2+}, where the K^+-stabilized dimer again dissociates at ≈80°C, but the Na^+-stabilized dimer dissociates at ≈40°C (data not shown). These experiments demonstrate that tightly bound K^+ must be considered an integral part of the HIV-1 dimerization domain under physiological conditions, since intracellular K^+ concentrations are >100 mM.

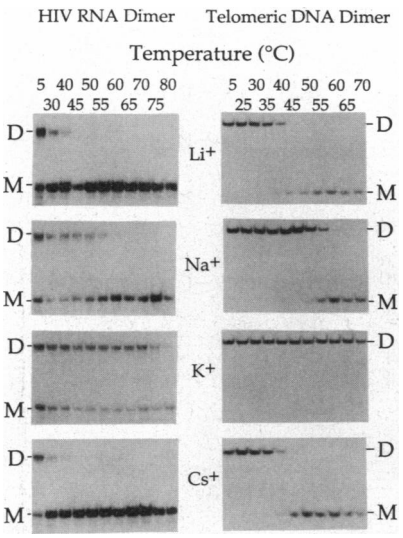


FIG. 3. Cation-dependent thermal stability of the HIV-1 RNA dimer. Thermal stabilities of the RNA 1 dimer (*Left*) and a control quadruple helical guanine hairpin dimer formed by the telomeric DNA sequence d(GTCGACCCGGGTTGGGGTTGGGGTTG-GGGTTGGGGTTGGGG)·d(CCCCAACCCCAACCCCAACCCGG-GTCGAC) (*Right*) were measured in parallel in buffers containing monovalent cations of different ionic radius. Preformed dimers were incubated in buffered solutions (10 mM Tris·HCl, pH 8.3) of 0.1 M LiCl, NaCl, KCl, or CsCl and, in the case of RNA 1, 5 mM $MgCl_2$. Aliquots were removed after 5-min incubations at each temperature, and dimer dissociation was assayed by nondenaturing PAGE. The experiment reveals size-specific K^+ binding within the dimerization domain of both the RNA 1 dimer and the telomeric DNA dimer.

Guanine-rich oligonucleotides derived from 5S RNA sequences also form intermolecular guanine quadruple helical structures *in vitro* (44). These tetrameric structures are again specifically stabilized by K^+ but are composed of four strands arranged in a parallel orientation as has been postulated for guanine- and inosine-containing RNA polymers (reviewed in ref. 23). Like their DNA analogues, RNA oligonucleotides tetramerize very slowly, requiring high oligonucleotide strand concentrations (>100 μM) and long incubation periods. In contrast, RNA 1 dimerizes at low temperature (37°C) and RNA concentrations (1.5 μM), although dimerization is still much slower than most other nucleic acid annealing processes ($k \approx 50$ liter/mol·sec).

The cation dependence of HIV-1 RNA dimerization has been examined previously, and the yield of dimer was shown to correlate inversely with the size of monovalent cation (i.e., $Li^+ > Na^+ > K^+ > Cs^+$) (22). Therefore, it was argued that G-tetrad formation alone could not account for the observed cation dependence of HIV-1 RNA dimerization. However, our thermal dissociation experiments are entirely consistent and even diagnostic of G-tetrad formation (compare Fig. 3 *Right* with Fig. 3 *Left*). This apparent paradox is resolved by considering that the rates of intermolecular G-tetrad formation and dissociation are both generally slow (24). Experiments that measure the yield of HIV-1 RNA dimer under different conditions therefore reflect the initial rates of dimer formation, whereas our thermal dissociation experiments reflect the rate of dimer dissociation (and/or overall thermodynamic stability) as a function of temperature. There is no reason to expect that the kinetics of dimer formation will mirror overall thermodynamic stability. Indeed, telomeric DNA dimerizes more rapidly in the presence of Na^+ than K^+ in spite of the fact that the dimer, once formed, has a higher affinity for K^+ than Na^+ (34). Furthermore, we observe empirically that model RNA sequences containing only guanine nucleotides exhibit the same inverse correlation between rate of dimer formation and cation size (e.g., $Li^+ > Na^+ > K^+$) and the same cation-dependent thermal stability (e.g., $K^+ > Na^+ > Li^+ \approx Cs^+$) as the HIV-1 RNA dimers (W.I.S. and D. P. Bogenhagen, unpublished data). Therefore, we conclude that the dramatically enhanced thermal stability of the HIV-1 RNA dimer in the presence of K^+ is strong evidence for intermolecular G-tetrad formation.

Model for the Dimerization of HIV-1 RNA. In the simplest possible model, dimerization requires that both RNA strands contribute two tracts of guanine nucleotides that are folded back on themselves to provide the four strands of a local quadruple helix. Deletion analyses at the 3′ end of the HIV-1 RNA dimerization domain indicated that the conserved purine-rich sequence GGGGGAGAA at positions 817–825 is required for dimerization and therefore presumably contributes one of the two tracts of guanine nucleotides required for guanine hairpin dimer formation. We anticipate that at least two other guanine bases must be contributed by a second tract elsewhere in the dimerization domain in order to create a K^+ binding site composed of stacked G-tetrads (see Fig. 1 *Lower*).

It is not clear, however, that a quadruple helical dimerization domain need necessarily be composed entirely of guanine nucleotides, since there is some evidence for quadruple helix formation by the mixed homopurine sequences poly(AAG), poly(AG), and poly(GGA) (where the A_4 tetrad is presumed to be isostructural with the G_4 tetrad shown in Fig. 1 *Lower*) (45). Of particular interest is the putative quadruple helical structure formed by poly(GGA), which binds K^+ in a size-selective fashion. However, the structure(s) formed by poly(GGA) remain to be characterized fully, since oligodeoxynucleotides composed of mixed purine sequences can also form parallel double helices (that do not exhibit anomalous monovalent cation-dependent stability) (46). Dimerization via a mixed G/A quadruple helical structure is consistent with sequences found in the region of retroviral dimerization signals, which frequently contain multiple runs of mixed adenosines and guanosines. In particular, the prevalence of the sequence motif Pu-G-G-A-Pu-A (Pu = purine) within the dimerization domains of a number of different retroviruses has been noted, and a role for this sequence motif in dimerization has been postulated (22). We note that the purine-rich sequence GG*GGGAGA*A at positions 817–825 that is required for RNA dimerization *in vitro* contains such a sequence (italic), in good agreement with this proposal.

It is likely that other structural elements within RNA 1 contribute to the rate and specificity of DLS formation. A model for the secondary structure of the monomeric HIV-1 RNA encompassing the 5′ end of the *gag* gene (630–840) has been proposed on the basis of computer folding, biochemical, and phylogenetic data (47). These experiments indicate that there are two stable stem–loop structures within the 5′ untranslated region of RNA 1 (742–755 and 760–783) but very little stable secondary or tertiary structure within the coding sequence of the *gag* gene itself (790–840). On the basis of this secondary structure as well as biochemical and genetic deletion experiments (15, 21, 48), the sequence GGAGG at positions 771–775 appears to be an attractive candidate for a second purine-rich region involved in dimerization. This sequence is exposed in the loop of a stem–loop structure in the model of the folded monomeric RNA (47), and chemical crosslinking studies indicate that nucleotides in the stem may compose part of a high-affinity binding site for the HIV-1 nucleocapsid protein (21). Furthermore, deletion of a 30-nt fragment (corresponding to 758–786 in HIV-1_{NL4-3}) encompassing this sequence has been shown to block dimerization in the context of a larger HIV-1_{Mal} RNA fragment (21). Other potential candidate purine-rich sequences include AGAAG-GAGAGAGA at 778–790 and GGG at 792–794, although it is

alternatively possible that a quadruple helix is composed of discontinuous purine nucleotides.

Biological Implications. It has been assumed that the RNA dimer that forms spontaneously *in vitro* is identical to the dimer linkage structure within the HIV-1 virion (21, 22). RNA dimers formed *in vitro* and *in vivo* appear similar by electron microscopy and have similar thermal stabilities (21). Our experiments indicate that guanine (and/or purine) base tetrads mediate the self-association of HIV-1 RNA *in vitro*. Therefore, we suggest that G-tetrads also mediate the alignment of HIV-1 RNA *in vivo*. By extension, it seems reasonable to expect that quadruple helical structures also serve to stabilize genomic RNA dimers in other retroviruses since the mechanisms of retroviral RNA dimerization appear to be conserved (21). It remains possible, however, that the high ionic strength conditions supporting dimerization of the HIV-1 RNA fragments *in vitro* fortuitously allow dimerization via a nonnative structure. Characterization of the RNA 1 dimer described herein should facilitate direct comparison of the two structures.

The ability of homologous retroviral RNA strands to spontaneously self-associate demonstrates that RNA itself can encode the potential for self-recognition required for genomic organization and recombination. It appears, however, that the rate of HIV-1 RNA dimerization observed *in vitro*, even under optimal conditions, is too slow to function on the time scale of a viral infection. Furthermore, once formed, dilute solutions of HIV-1 RNA dimers are stable for hours under physiological conditions (W.I.S., unpublished data). Therefore, it appears that the biophysical problem inherent in using G-tetrads in biological systems is kinetic rather than thermodynamic. Darlix and coworkers have shown, however, that the highly basic HIV-1 nucleocapsid protein (encoded by the 3′ end of the *gag* gene) facilitates formation of an HIV-1 RNA dimer that is indistinguishable in stability and electrophoretic mobility from the RNA dimer formed spontaneously in the absence of protein (21). The nucleocapsid protein greatly enhances the rate of RNA dimerization and alleviates the need for high salt incubation conditions *in vitro*. Furthermore, genetic evidence is consistent with a role for the nucleocapsid protein in HIV-1 RNA dimerization *in vivo* (15).

In an increasing number of nucleic acid–protein systems, it is becoming apparent that the nucleic acid component contains the essential information for recognition and/or catalysis, while the protein cofactor simply allows the nucleic acid to perform its function more rapidly, with greater specificity, and under conditions of lower ionic strength. Examples of such systems include RNase P (49) and possibly also peptidyl transferase (50), mRNA splicing (reviewed in ref. 51), and Rec A-mediated DNA recombination (52). We speculate that the nucleocapsid protein–HIV-1 RNA complex may also fall into this category, with the potential for self-recognition encoded within the RNA sequence and the nucleocapsid protein functioning as a kind of RNA chaperone (albeit noncatalytic) to lower the activation barrier for quadruplex formation. This would allow HIV-1 RNA genomes to associate rapidly under physiological conditions as required by the viral life cycle.

This work was initiated when both authors were members of the Medical Research Council Laboratory of Molecular Biology. We thank our former colleague Dr. D. Brown for initially suggesting to us possible analogies between the dimerization of telomeric DNA and HIV-1 RNA. We also thank Drs. D. Brown and A. Klug for their helpful suggestions and continued support and Drs. D. Carroll, B. Bass, A. Christensen, and P. Kosa for critical reading of the manuscript. S.H. is a Medical Research Council Senior AIDS Research Fellow.

1. Linial, M. L. & Miller, A. D. (1990) *Curr. Top. Microbiol. Immunol.* **157,** 124–152.
2. Stuhlmann, H. & Berg, P. (1992) *J. Virol.* **66,** 2378–2388.
3. Coffin, J. M. (1979) *J. Gen. Virol.* **42,** 1–26.
4. Junghans, R. P., Boone, L. R. & Skalka, A. M. (1982) *Cell* **30,** 53–62.
5. Kung, H.-J., Hu, S., Bender, W., Bailey, J. M., Davidson, N., Nicolson, M. O. & McAllister, R. M. (1976) *Cell* **7,** 609–620.
6. Luban, J. & Goff, S. P. (1991) *J. Virol.* **65,** 3203–3212.
7. Bieth, E., Gabus, C. & Darlix, J.-L. (1990) *Nucleic Acids Res.* **18,** 119–127.
8. Prats, A.-C., Sarih, L., Gabus, C., Litvak, S., Keith, G. & Darlix, J.-L. (1988) *EMBO J.* **7,** 1777–1783.
9. Prats, A.-C., Roy, C., Wang, P., Erard, M., Housset, V., Gabus, C., Paoletti, C. & Darlix, J.-L. (1990) *J. Virol.* **64,** 774–783.
10. Katoh, I., Kyushiki, H., Sakamoto, Y., Ikawa, Y. & Yoshinaka, Y. (1991) *J. Virol.* **65,** 6845–6855.
11. Prats, A.-C., Housset, V., de Billy, G., Cornille, F., Prats, H., Roques, B. & Darlix, J.-L. (1991) *Nucleic Acids Res.* **19,** 3533–3541.
12. Méric, C., Darlix, J.-L. & Spahr, P.-F. (1984) *J. Mol. Biol.* **173,** 531–538.
13. Méric, C. & Spahr, P.-F. (1986) *J. Virol.* **60,** 450–459.
14. Méric, C. & Goff, S. P. (1989) *J. Virol.* **63,** 1558–1568.
15. Aldovini, A. & Young, R. A. (1990) *J. Virol.* **64,** 1920–1926.
16. Gorelick, R. J., Henderson, L. E., Hanser, J. P. & Rein, A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8420–8424.
17. Fu, X., Katz, R. A., Skalka, A. M. & Leis, J. (1988) *J. Biol. Chem.* **263,** 2140–2145.
18. Anderson, D. J., Lee, P., Levine, K. L., Sang, J., Shah, S. A., Yang, O. O., Shank, P. R. & Linial, M. L. (1992) *J. Virol.* **66,** 204–216.
19. Duesberg, P. H. (1968) *Proc. Natl. Acad. Sci. USA* **60,** 1511–1518.
20. Mangel, W. F., Delius, H. & Duesberg, P. H. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 4541–4545.
21. Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F. & Barré-Sinoussi, F. (1990) *J. Mol. Biol.* **216,** 689–699.
22. Marquet, R., Baudin, F., Gabus, C., Darlix, J.-L., Mougel, M., Ehresmann, C. & Ehresmann, B. (1991) *Nucleic Acids Res.* **19,** 2349–2357.
23. Guschlbauer, W., Chantot, J.-F. & Thiele, D. (1990) *J. Biomol. Struct. Dyn.* **3,** 491–511.
24. Sundquist, W. I. (1991) *Nucleic Acids Mol. Biol.* **5,** 1–24.
25. Sen, D. & Gilbert, W. (1991) *Curr. Opin. Struct. Biol.* **1,** 435–438.
26. Sundquist, W. I. & Klug, A. (1989) *Nature (London)* **342,** 825–829.
27. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. (1989) *Cell* **59,** 871–880.
28. Sen, D. & Gilbert, W. (1990) *Nature (London)* **344,** 410–414.
29. Panyutin, I. G., Kovalsky, O. I. & Budowsky, E. I. (1989) *Nucleic Acids Res.* **17,** 8257–8271.
30. Panyutin, I. G., Kovalsky, O. I., Budowsky, E. I., Dickerson, R. E., Rikhirev, M. E. & Lipanov, A. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 867–870.
31. Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992) *Nature (London)* **356,** 126–131.
32. Zimmerman, S. B., Cohen, G. H. & Davies, D. R. (1975) *J. Mol. Biol.* **92,** 181–192.
33. Pinnavaia, T. J., Marshall, C. L., Mettler, C. M., Fisk, C. L., Miles, H. T. & Becker, E. D. (1978) *J. Am. Chem. Soc.* **100,** 3625–3627.
34. Oka, Y. & Thomas, C. A., Jr. (1987) *Nucleic Acids Res.* **15,** 8877–8898.
35. Raghuraman, M. K. & Cech, T. R. (1990) *Nucleic Acids Res.* **18,** 4543–4552.
36. Smith, F. W. & Feigon, J. (1992) *Nature (London)* **356,** 164–168.
37. Blackburn, E. H. (1991) *Nature (London)* **350,** 569–573.
38. Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. (1991) *Nature (London)* **350,** 718–720.
39. Milligan, J. F. & Uhlenbeck, O. C. (1989) *Methods Enzymol.* **180,** 51–62.
40. Adachi, A., Gendelman, H. E., Koenig, S., Folks, T., Willey, R., Rabson, A. & Martin, M. A. (1986) *J. Virol.* **59,** 284–292.
41. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Plainview, NY).
42. Fasman, G. D., ed. (1975) *Handbook of Biochemistry and Molecular Biology*, (CRC, Cleveland), 3rd Ed., Vol. 1, p. 589.
43. Myers, G., Berzofsky, J. A., Korber, B. & Smith, R. F., eds. (1991) *Human Retroviruses and AIDS* (Los Alamos Natl. Lab., Los Alamos, NM).
44. Kim, J., Cheong, C. & Moore, P. B. (1991) *Nature (London)* **351,** 331–332.
45. Lee, J. S. (1990) *Nucleic Acids Res.* **18,** 6057–6060.
46. Rippe, K., Fritsch, V., Westhof, E. & Jovin, T. M. (1992) *EMBO J.* **11,** 3777–3786.
47. Harrison, G. P. & Lever, A. M. L. (1992) *J. Virol.* **66,** 4144–4153.
48. Lever, A., Gottlinger, H., Haseltine, W. & Sodroski, J. (1989) *J. Virol.* **63,** 4085–4087.
49. Reich, C., Olsen, G. J., Pace, B. & Pace, N. R. (1988) *Science* **239,** 178–181.
50. Noller, H. F., Hoffarth, V. & Zimniak, L. (1992) *Science* **256,** 1416–1419.
51. Steitz, J. A. (1992) *Science* **257,** 888–889.
52. Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. (1990) *Genes Dev.* **4,** 1951–1963.