3-31-1999

Monomer-Dimer Equilibrium Constants of RNA in the Dimer Initiation Site of Human Immunodeficiency Virus Type 1

Michael F. Shubsda
Syracuse University

Mark P. McPike
Syracuse University

Jerry Goodisman
Syracuse University

James C. Dabrowiak
Syracuse University

Recommended Citation

Shubsda, M. F., McPike, M. P., Goodisman, J., & Dabrowiak, J. C. (1999). Monomer - dimer equilibrium constants of RNA in the dimer initiation site of human immunodeficiency virus type 1. Biochemistry, 38(31), 10147-10157

Monomer−Dimer Equilibrium Constants of RNA in the Dimer Initiation Site of Human Immunodeficiency Virus Type 1†

Michael F. Shubsda, Mark P. McPike, Jerry Goodisman, and James C. Dabrowiak*

Department of Chemistry, Center for Science and Technology, R 1-014, Syracuse University, Syracuse, New York 13224-4100

Received March 31, 1999; Revised Manuscript Received May 19, 1999

ABSTRACT: The genome of the human immunodeficiency virus (HIV) exists as a dimer of two identical RNA molecules hydrogen bonded to each other near their 5′ ends. The dimer, known to be important for viral infectivity, is formed by two monomers interacting through a stem−loop structure called the dimer initiation site (DIS). An initially formed intermediate, the "kissing" dimer, is unstable and rearranges to the stable, duplex form. In this report we use nondenaturing polyacrylamide gel electrophoresis to measure the monomer−dimer equilibrium constant of three RNA sequences, 41-, 27-, and 19-mers, located in the DIS of the MAL isolate of HIV-1. Experiments in which the RNA was equilibrated at various temperatures before electrophoresis revealed that interconversion is rapid for all the sequences, so that they reach equilibrium in the loading well of the gel at 5 °C before they enter the gel proper. However, interconversion kinetics in the gel are slow, so autoradiographic spot intensities can be used to measure the amounts of monomer and dimer present when the sample entered the gel. After correction for the amount of RNA added with the radiolabel and dilution of samples in the loading well of the gel, dimerization equilibrium constants were calculated from spot intensities. The calculated values of the dimerization constant K at 5 °C were $\sim 10^5$, $\sim 10^6$, and $\sim 10^8$ M^{-1} for the 41-, 27-, and 19-mers, respectively, in solutions of ionic strength, I, of about 100 mM. The decrease in K by three orders of magnitude between the 19-mer and 41-mer is due in part to the change in rotational entropy of rodlike molecules on dimerization and in part to the increased conformational entropy of the monomers. As expected, increased ionic strength increases the dimerization constant for all three RNAs. For the 41-mer, however, K has a maximum value at I ~140 mM. The origin of the decrease in K for higher I is unknown but it may be due to formation of species (perhaps higher order oligomers) that do not enter the gel. The 41-mer exists in two dimeric forms assigned to the kissing and duplex dimers. The ratio of kissing to duplex form at 5 °C is 0.48 ± 0.22 at I = 113 mM and 0.91 ± 0.35 at I = 183 mM. The observed decrease in K with RNA length suggests that the dimerization constant of the packaging region of HIV-1 is small, $< \sim 10^5$ M^{-1}, implying that the nucleocapsid protein is important in promoting dimerization in the capsid of the virus.

Two copies of viral RNA are incorporated as a dimer into the capsid of the human immunodeficiency virus type 1 (HIV-1) virion as it forms (*1*). The region of dimerization, referred to as the ψ-element or packaging region, is important for a number of viral functions including encapsidation and translation of the genome, reverse transcription, and stabilization of the genome against degradation (ref *2* and references therein). In the monomeric form of the genome, the ψ-element is thought to exist as four stem−loop-type structures (*2*−*9*) (Figure 1). One of the loops (SL1) that is part of the dimer initiation site (DIS) contains a 6-base palindrome GCGCGC or GUGCAC, flanked by one 3′ and one or two 5′-adenosines (*10*). When the dimer forms from the monomer, the SL1 portions of two monomers are believed to combine to form a "kissing loop" dimer (*11*−*13*). This intermediate, which has the palindromes from two monomers hydrogen-bonded to each other, is unstable and it ultimately rearranges to a stable dimer by converting the *intra*molecular Watson−Crick base pairing associated with each monomer into *inter*molecular base pairing between two monomers (*9*). The formation of the kissing dimer is facilitated by the binding of the zinc finger protein NCp7, a segment of the structural gag protein, to RNA (*13*). This drives the equilibrium to the extended (duplex) dimer, ultimately producing the "packaged" diploid form of the genome in the mature virus particle.

The solution structure of the kissing dimer has recently been characterized by NMR. A 23-base RNA including the loop and two distal stem base pairs from SL1 (LAI strain) was subjected to conditions that favored the formation of the kissing loop structure (*14*). The NMR analysis showed that the kissing dimer is stabilized by the palindromic sequence in the loop, GCGCGC, that forms Watson/Crick base pairs, and adenine bases flanking the palindrome, that enter into a triple base array with the opposing strand. Dardel et el. (*15*) also used NMR to characterize a kissing dimer that forms between two 19-mer RNAs from the MAL strain of the virus. In this case the palindromic sequence, GUGCAC, is involved in Watson/Crick base pairs and the two adenines that flank the sequence, one on the 3′ and one on the 5′ side, are stacked inside of the helix forming an intramolecular noncanonical A·A base pair.

† This work was supported in part by NIH Grant GM32691 to P. N. Borer.

* Corresponding author: phone (315) 443-4601; Fax (315) 443-4070; e-mail, jcdabrow@syr.edu.

10.1021/bi990744t CCC: $18.00
Published on Web 07/17/1999

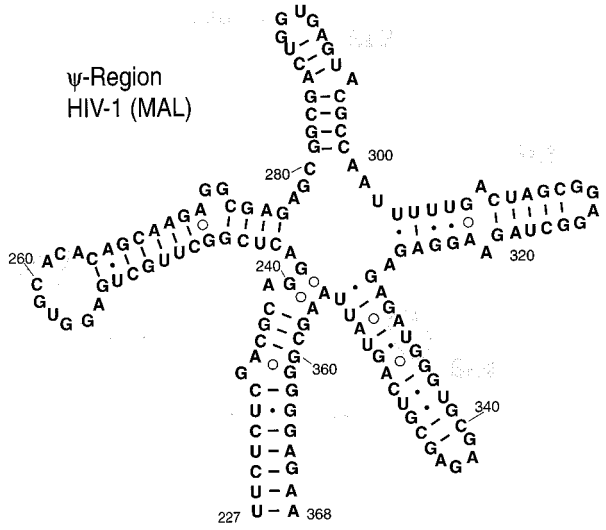

FIGURE 1: Packaging region (ψ-element) of HIV-1 RNA (MAL), showing known functional regions of the ψ-element and probable secondary structure of the monomeric form including presumptive G○A and U•G base pairs (*9*). Base numbering is from the full-length mRNA. Shown are stem loops SL1−4 and the stem connecting the packaging region to the remainder of the genome, sA and sA*. DIS = dimer initiation site, SD = 5′ splice donor site, and gag = start codon for the *gag* gene.

The importance of the HIV dimer for viral infectivity and the potential of the dimer initiation site as a target for drug action (*16*) prompted us to measure the monomer/dimer equilibrium constant, *K*, of the HIV-1 RNA sequences shown in Figure 2. From the magnitude of the equilibrium constant one can determine the free energy of interaction, an important quantity for designing new types of RNA-specific drugs capable of blocking dimer formation and reducing viral infectivity. Previously, we used nondenaturing gel electrophoresis to measure the monomer−dimer equilibrium constants and study the kinetics of small DNA molecules (*17*). As an aid to interpretation of experiments, the differential equations associated with diffusion, migration, and chemical reaction were solved, and intensity patterns were generated and compared with experimental results from electrophoresis measurements. In this report we use nondenaturing gel electrophoresis to measure monomer−dimer equilibria of RNA sequences from the dimer initiation site of HIV-1 Mal, Figure 2.

MATERIALS AND METHODS

Synthesis of RNA. The HIV RNAs, having the sequences 5′-GGACUCGGCUUGCUGAGGUGCACACAG-CAAGAGGCGAGUCC-3′ (41-mer), 5′-GGCUUGCUGAG-GUGCACACAGCAAGAC-3′ (27-mer), and 5′-GGCUGAG-GUGCACACAGCC-3′ (19-mer) were made by transcribing a DNA template (Midland) containing the appropriate coding and promoter sequences using T7 polymerase in the presence of the four ribonucleotide triphosphates. To maximize the yield of RNA in the synthesis and produce structures having minimal fraying in the stems, the normal coding sequences in the DIS of HIV-1 Mal were slightly modified at the ends of the stems.

Trial transcription reactions to make RNA were first carried out in 50 μL volumes with 200 nM template strand, 300 nM of the 17-nucleotide long T7 promoter strand (*18*), 2 mM NTP, 8.1 mM $MgCl_2$, ∼60 units of T7 polymerase in the transcription buffer (40 mM Tris-HCl, 1 mM spermidine, 0.01% Triton X-100, and 5 mM DTT, pH 8.1), at 37 °C for 3.5 h (*19*). Portions of the trial reaction mixtures were separated by denaturing PAGE and the products were visualized with Stainsall (Sigma) to determine the approximate yield of full-length oligomer. The trial reaction conditions were optimized by changing the amounts of $MgCl_2$, NTP, and enzyme to maximize the yield of full length oligomer. Once transcription conditions were optimized, the reaction was scaled to a total volume of 5 mL and the RNA was purified by denaturing PAGE. With the help of UV backshadowing, the band containing the full-length oligomer was excised and the RNA was recovered from the gel slab by electroelution in 1 × TBE buffer (90 mM Tris−borate and 1 mM EDTA, pH 8) (*20*, *21*). The RNA was desalted with a Centricon concentrator.

The 5′-phosphate of the RNA was removed with calf intestinal phosphatase, CIP (Boehringer Mannheim), in 60 μL of CIP buffer containing ∼5 μM RNA and 1 unit of CIP (*20*). After 1 h at 50 °C, the enzyme was removed by extraction with phenol and the RNA was recovered by precipitation with cold ethanol. The RNA was 5′-end labeled in the presence of [γ-^{32}P]ATP (Amersham) and T4 polynucleotide kinase (Promega). The labeling solution contained 15 mM RNA, 50 μCi of [γ-^{32}P]ATP and 20 units of polynucleotide kinase in a total volume of 20 μL in kinase reaction buffer (*20*). After incubation for 1 h at 37 °C, the labeled oligomer was precipitated by addition of cold ethanol and purified by 20% denaturing PAGE with 1× TBE running buffer. After visualization of the labeled RNA by autoradiography, the portion of the gel containing the oligomer was excised and the labeled RNA was recovered by suspending the crushed gel segment in 1 mL of 1× TBE buffer (*20*). A 5 μL sample of oligomer recovered in this manner yielded 60 000 cpm on scintillation counting.

To determine whether the various forms observed under nondenaturing PAGE conditions were in fact due to a single-length oligomer, the gel-purified, end-labeled RNA molecules were electrophoresed a second time under denaturing PAGE conditions. In all cases, autoradiography showed only one band, indicating that the labeled RNA molecules were homogeneous with regard to length. To show that the fastest running band in the native gels is the monomer, serial dilutions to ∼0.01 μM of only the radiolabeled RNA were carried out and analyzed by native PAGE. In all cases only one band, the monomer in the stem−loop form, was observed in these experiments.

Determination of the Molar Extinction Coefficients of RNA. Accurate molar extinction coefficients at 260 nm, ϵ_{260}, for the monomeric forms of the RNA molecules were determined by measuring the absorbance of the intact RNA at ∼1 μM, and the absorbance after complete hydrolysis in base (5 M NaOH, 25 °C, 1 day), according to previously published methods (*22*, *23*). The hydrolysate was brought to neutral pH by addition of 5 M HCl/1 M Na_2HPO_4 and then the absorbance was measured a second time at 260 nm. The values of ϵ_{260} for the intact RNAs were calculated from the known molar extinction coefficients of the four ribonucleotides and the number of each nucleotide present in the various sequences. The measured molar extinction coefficients (M^{-1} cm^{-1}) were as follows: 41-mer (two determinations,

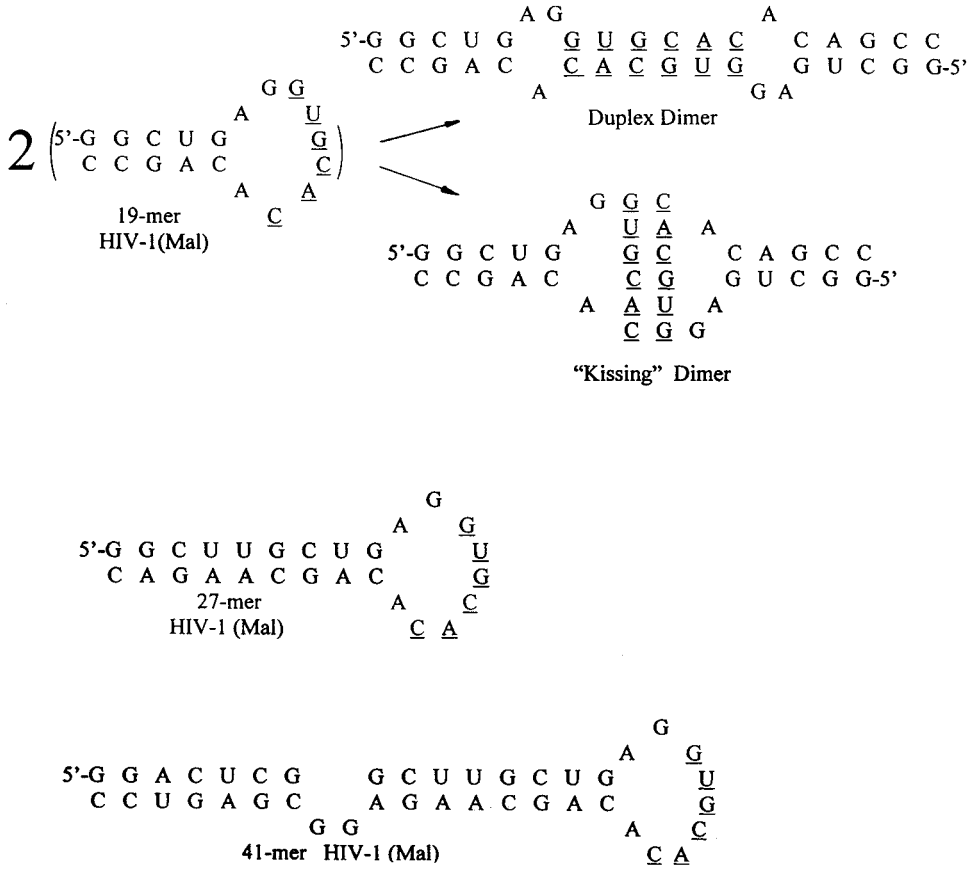

FIGURE 2: Sequences of RNA from the dimer initiation site located in SL1 of HIV-1 (MAL) used in the study. The monomers are shown as stem−loop (hairpin) structures. Dimers involving all three sequences can exist in either a kissing or duplex form. This is shown for the 19-mer in the figure.

$(3.32 \pm 0.1) \times 10^5$; 27-mer (two determinations), $(1.98 \pm 0.08) \times 10^5$; 19-mer (three determinations), $(1.49 \pm 0.02) \times 10^5$.

Monomer−Dimer Equilibria. Most experiments were carried out in the reaction buffer, 1× or 0.5× TBE. Preliminary studies were performed in which 41-mer RNA was heat-denatured at 100 °C for 10 min and then rapidly cooled on ice, followed by incubation for 2−24 h at temperatures from 5 to 65 °C before being loaded into the gel. The resulting intensity patterns in the gel were *independent* of the incubation conditions outside of the gel, indicating that the RNA comes to equilibrium at the gel temperature during the 10 or 15 min it remains in the gel well, before entering the gel proper. Therefore, all subsequent samples were equilibrated at 5 °C for 2 h after heat denaturation at 100 °C for 10 min. To examine the effects of salt on the equilibrium, various concentrations of NaCl were added after the heat denaturing step. After the sample was equilibrated, a 6 μL aliquot of the reaction mixture was mixed with 2 μL of a 50% glycerol/water mixture (at 5 °C) and vortexed briefly.

Then 5 μL of the mixture was loaded on a nondenaturing polyacrylamide gel, at 5 °C, with the current flowing to minimize dead time in the well. Electrophoresis was carried out (running buffer, 1× TBE) at 1500 V for 11−15 h using a standard sequencing apparatus (Sigma) in a refrigerator. The temperature of the buffer wells (±2 °C) was the temperature reported for the run.

Dilution in the Well of the Gel. During the time a loaded sample remains in the gel well, RNA and components of the loading buffer diffuse out of the glycerol sample and mix with electrophoresis running buffer. (This is shown in experiments with dye dissolved in the glycerol.) For rapidly interconverting systems such as ours, the equilibrium state is changed to one associated with the new concentrations in the well. In our experiments, a volume V_i (usually 5 μL) of sample was loaded into the bottom of the well, which had a volume of V_w and contained 1× TBE electrophoresis running buffer. Earlier work with DNA (*17*) showed that the effect of mixing in the well could be estimated by lowering the concentration of the oligonucleotide by a factor $d = V_i/V_w$ and changing the ionic strength of the solution from I_i to $dI_i + (1 - d)I_w$, where I_w is the ionic strength of the electrophoresis running buffer. The ionic strength of the 1× TBE running buffer is 0.102 M. An approximation to V_w was obtained by multiplying the dimensions of the tooth used to make the well, which gave 21.6 μL.

The dilution of the sample in the well by a factor of d means that the equilibrium mixture in the gel well has a total RNA concentration (labeled + unlabeled) of C_0d, where C_0 is the original concentration. The diluted concentrations of RNA, $C = C_0d$, were used in deriving the equilibrium constant, K, from KC, calculated according to eq 4 below from the spot intensities on the autoradiogram.

Quantitation of Autoradiographic Data. Autoradiography was done at −20 °C with Kodak X-Omat film. Typically, more than one exposure of a gel was required to obtain measurable images for which most optical densities were in the linear response range of the film, optical density (OD) less than about 1.1. The autoradiogram was scanned to a

JPEG or TIF file with a Hewlett-Packard Scanjet 5P scanner. The scanned images were analyzed with Sigma Scan software (Jandel Scientific) to yield plots of optical density as a function of position (electrophoresis distance) for each lane of the autoradiogram with an 11-pixel wide line scan. (The lanes themselves were about 33 pixels wide.) The resulting profile was fitted to a sum of Gaussians and a linear background. The same number of Gaussians was used for all lanes for a given oligomer. When a peak exhibited saturation at its maximum (OD > 1.1), data points in the region of saturation were eliminated prior to curve fitting and analysis. The resulting integrated intensities were comparable to those obtained from short autoradiographic exposure times. Intensities of peaks reported are the calculated areas of the Gaussians, in arbitrary units (average optical density multiplied by number of pixels).

THEORY

Monomer/Dimer Equilibrium Constants from Gel Data. Let the concentrations of monomeric (stem−loop) and dimeric (kissing, duplex) RNA be [M] and [D], respectively. The equilibrium constant for the dimerization reaction

$$M + M \rightleftharpoons D \tag{1}$$

is written:

$$K = \frac{[D]}{[M]^2} \tag{2}$$

To obtain the concentrations of monomeric and dimeric RNA from the intensities of the corresponding spots on a nondenaturing polyacrylamide gel, one first notes that the intensities come from radiolabeled RNA, which constitutes a small fraction, say *f*, of the total oligomer present. The ratio of the concentration of radiolabeled monomer to the total concentration of monomer is *f*, the ratio of the concentration of singly labeled dimer to the total concentration of dimer is *2f*, and the ratio of doubly labeled dimer to the total is f^2. Since *f* is small, all the labeled dimer may be considered to have a single label, and f^2 may be neglected compared to *2f*. The measured intensity being proportional to the concentration of radiolabeled RNA, the intensity of the monomer spot is $\theta_M = cf[M]$ and the intensity of the dimer spot is $\theta_D = 2cf[D]$, where *c* is the unknown constant of proportionality. The equilibrium constant is

$$K = \frac{1}{2}\frac{\theta_D/cf}{(\theta_M/cf)^2} \tag{3}$$

Since the value of *c* is not known, we multiply both sides of the equation by *C*, the total concentration of oligonucleotides, expressed as monomers; *C* is equal to θ_T/cf, where θ_T is the total intensity, $\theta_M + \theta_D$. Now

$$KC = \frac{1}{2}\frac{\theta_D\theta_T}{(\theta_M)^2} \equiv Q \tag{4}$$

which does not depend on *c* or *f* and involves measurable spot intensities only. (These spot intensities are peak areas in a plot of intensity vs position in the electrophoresis direction.) Another reason to use eq 4 is that the right-hand side is independent of the amount of RNA, so it automatically "normalizes," correcting for variations in the amount of oligomer loaded into the various lanes of the gel.

Electrophoresis causes the monomer and dimer to separate and perhaps reequilibrate. Only if monomer−dimer interconversion in the gel is unimportant will band intensities for the monomer and dimer in the gel accurately reflect the amounts of the two forms present in solution before electrophoresis. Following earlier workers (*24*, *25*), we calculated intensity patterns by solving the differential equations governing diffusion, migration, and interconversion of the monomer and dimer (*17*). It was found that, depending on the rate of monomer−dimer equilibration in the gel, the intensity patterns can have from one to four features.

When reaction rates for monomer−dimer conversion in the gel are very rapid, the pattern consists of a single peak, representing an interconverting mixture of monomer and dimer. Relative amounts of monomer and dimer cannot be inferred from this pattern without additional information. When reaction rates for conversion in the gel are very slow, the intensity pattern consists of two peaks, for monomer and dimer. In this case, relative amounts of monomer and dimer are obtainable without ambiguity and give the relative amounts in solution before electrophoresis. For higher conversion rates, one sees, in addition to monomer and dimer peaks, streaks or shoulders extending from one peak in the direction of the other; these represent monomer formed from dimer, and dimer formed from monomer, in the gel. If intense enough, two such streaks may combine into a broad peak between the monomer and dimer peaks, giving a three-peak pattern.

When the presence of features other than two peaks shows that monomer−dimer interconversion in the gel is important, eq 4 does not give the value of *KC* for equilibrium in homogeneous solution. Although one may try to assign the streaks between the monomer and dimer peaks to monomer and dimer, it is better to perform a load-and-run experiment, in which identical samples are loaded into different lanes of the gel at different times, while the electrophoresis is proceeding. Then monomer and dimer intensities (or apparent values of *KC*) may be measured as a function of running time and extrapolated back to time 0, giving the values of intensities (or *KC*) at the time the molecules entered the gel. This was not necessary for these RNAs at 5 °C.

RESULTS AND ANALYSIS

In a number of experiments, the RNA was allowed to equilibrate in solution for various times and at various temperatures outside the gel before it was loaded into a nondenaturing gel for electrophoresis at 5 °C. For all RNAs the results were independent of equilibration temperature, showing that equilibrium was reestablished during the time the sample remained in the gel well at 5 °C. Subsequently, all samples were held at 5 °C before loading into the well of the gel. Experiments as a function of total RNA concentration and ionic strength *I* (adjusted by adding NaCl to the reaction medium) were then carried out. In all cases, there was little intensity outside the monomer and dimer spots, showing that reaction in the gel is slow relative to migration. It was observed long ago (*26*) that dissociation and recombination in the gel may be orders of magnitude slower than in bulk solution.

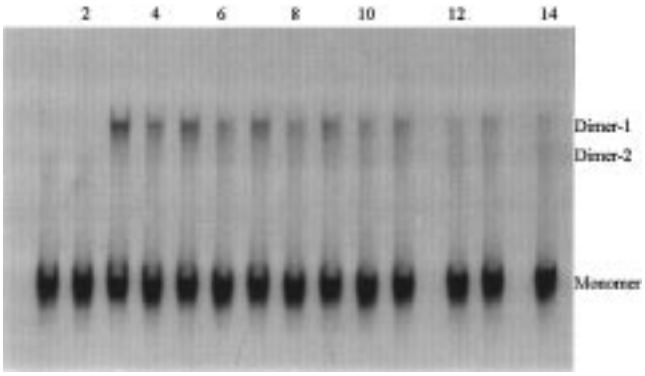

FIGURE 3: Autoradiogram of the 41-mer at 5 °C. Even- and odd-numbered lanes have ionic strengths $I = 113$ and 182 mM, respectively. The lane numbers and RNA concentrations (in micromolar, not including radiolabeled RNA) are as follows: lanes 1 and 2, 0 (radiolabeled RNA only); lanes 3 and 4, 1.89; lanes 5 and 6, 1.39; lanes 7 and 8, 1.04; lanes 9 and 10, 0.69; lanes 11 and 12, 0.35; lanes 13 and 14, 0.17.

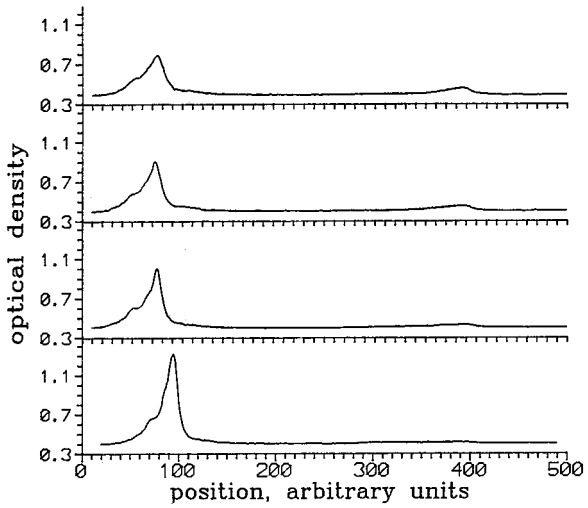

FIGURE 4: Scans from the autoradiogram (41-mer) shown in Figure 3. Ionic strength is 182 mM and [RNA] (not including radiolabeled RNA) = 0.17, 0.69, 1.39, and 1.89 μM, bottom to top.

Figure 3 shows the band pattern of the 41-mer as a function of total RNA concentration for two different values of ionic strength, $I = 113$ mM and $I = 182$ mM. From experiments incorporating molecular weight markers and the manner in which the bands vary with [RNA], it can be concluded that the faster running band is the monomer in the stem−loop form and the slower band is the dimer. Simulations for the reacting and migrating system (*17*) show that the low-intensity streak near the dimer band could be due to dimer formed from two monomers *during* electrophoresis or to a second form of the dimer. For the purposes of calculating K, the sum of the intensities of these bands was used as the intensity of the dimer. Scans of various lanes of the autoradiogram in Figure 3 are shown in Figure 4. To obtain measurable intensities for the dimer bands, it was necessary to have optical densities in the monomer band greater than 1.1, requiring the use of the curve-fitting analysis described in the Materials and Methods section.

We calculate a value of $Q = \theta_D(\theta_M + \theta_D)/2\theta_M^2$ for each RNA concentration C, according to eq 4, using for θ_M and θ_D the intensities or areas of the monomer and dimer bands. The area of the dimer band includes the streak area, as if it were a second form of the dimer. The values of Q obtained for the autoradiogram shown in Figure 3 are given in Table 1. If each value of Q is divided by C to get the apparent equilibrium constant K, one finds that the values decrease with C. This is because the RNA concentration is actually $C + L$, where L is the concentration of RNA added with the radiolabel, so the quantity Q actually represents $K(C + L)$. Therefore, a plot of Q vs C should be linear with slope equal to K and x-intercept equal to $-L$. The data of Table 1 are plotted in this way in Figure 5, with the best-fit lines given ($r^2 = 0.84$ and 0.90, respectively).

Table 1: Spot Intensities and Q[a] for 41-mer, Gel Temperature 5 °C

[RNA] (μM)	intensities ($I = 113$ mM)			intensities ($I = 183$ mM)		
	monomer	dimer	Q	monomer	dimer	calcd Q
0[b]	12.86	0	0	15.68	0	0
0.174	32.46	0.46	0.0072	20.32	1.26	0.0330
0.347	19.77	0.80	0.0211	18.83	1.67	0.0482
0.694	22.18	1.91	0.0468	15.20	2.05	0.0764
1.042	18.69	1.47	0.0424	13.85	2.47	0.1049
1.389	21.20	1.82	0.0465	14.04	2.20	0.0907
1.892	18.21	1.93	0.0586	12.70	2.81	0.1351

[a] $Q = \theta_D\theta_T/(2\ \theta_M^2)$, where θ_D, θ_M, and θ_T are dimer, monomer, and total spot intensities. [b] Radiolabeled RNA only.

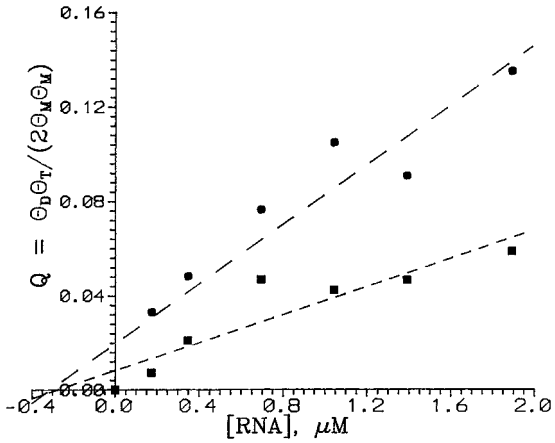

FIGURE 5: Plot of $Q = {}^1/_2\theta_D\theta_T/\theta_M^2$ vs RNA concentration for the 41-mer experiment shown in Figure 3 and Table 1.

From the linear least-squares fits we obtain $K = (0.0297 \pm 0.0059) \times 10^6\ M^{-1}$ and $(0.0632 \pm 0.0095) \times 10^6\ M^{-1}$ for ionic strengths $I = 113$ mM and $I = 183$ mM; values of L are (0.278 ± 0.207) μM and $(0.\ 312 \pm 0.159)$ μM, respectively. The stated errors are standard errors from the least-squares fit; the errors in the values of K are likely to be larger because they include errors in measurement of peak areas. In particular, small peak areas (less than 2) may be seriously underestimated and large-area peaks may be saturated (optical density greater than 1.1), making their areas hard to estimate.

If the slowest-running peak and the streak adjoining it actually represent two forms of a dimer, say A and B, respectively, the equilibrium constant for the conversion of one form to another would be

$$K' = [B]/[A] \tag{5}$$

The ratio of concentration is equal to the ratio of intensities R and should be independent of total RNA concentration. The linear least-squares fit of R versus RNA concentration

Table 2: Spot Intensities[a] and Calculated Q[b] for 41-mer

[RNA] (μM)	dimer	monomer	Q^b
0	0.99	25.56	0.020
0.005	1.79	31.77	0.030
0.009	1.44	29.26	0.026
0.023	1.88	32.54	0.031
0.039	1.76	26.46	0.035
0.093	1.87	30.65	0.032
0.235	3.57	30.99	0.064
0.391	4.37	30.71	0.081

[a] Temperature = 5 °C, ionic strength = 105 mM. [b] $Q = \theta_D\theta_T/(2\theta_M^2)$.

gives a slope of 3.5 ± 3.0, with a large error because one point is far from the line. Dropping this point gives the intensity ratio, which equals K', equal to ~1 (averaged over all ionic strengths). The ratio of the intensities of the two dimer spots on other gels (data not shown) indicates that $K' = 0.48 \pm 0.22$ at 113 mM ionic strength and $K' = 0.91 \pm 0.35$ at 183 mM ionic strength. It may also be noted that the total intensity (dimer plus monomer bands) is usually lower at $I = 183$ mM than at $I = 113$ mM.

Table 2 gives results from a second experiment involving the 41-mer; RNA concentrations from 0 to 0.39 μM were used, and the ionic strength in the well was 105 mM. The linear least-squares fit of Q vs C has $r^2 = 0.959$. From the slope of this line, $K = (0.147 \pm 0.012) \times 10^6$ M^{-1}; the negative of the x-intercept is $L = (0.172 + 0.044)$ μM. The value of K is several times larger than what was obtained from the data of Table 1, but this value is less accurate because of the large error in measuring the dimer intensity, which is small with the lower RNA concentrations used in this experiment.

Returning to Figure 5, a plot of Q vs [RNA], it appears that the data points at high concentration have a slope that is lower than the points at low concentration. This behavior has been observed for the results of other experiments with this oligomer (data not shown here). It may be associated with increased reaction in the gel or with formation of higher order complexes that do not enter the gel. Both effects should be larger at higher RNA concentrations. If only the first four points are used in linear fits, the data of Figure 5 give lines with $r^2 = 0.99$ and 0.95 for the two ionic strengths. The values of K are $(0.069 \pm 0.005) \times 10^6$ M^{-1} and $(0.105 \pm 0.016) \times 10^6$ M^{-1}, and the calculated values of L are (-0.033 ± 0.029) μM and (0.072 ± 0.063) μM, respectively. That these values of K, and not the smaller values originally calculated from Figure 5, are correct is shown by the results of experiments designed to measure K as a function of ionic strength.

Because the monomer is negatively charged, electrostatic screening, which increases with ionic strength, should favor dimerization and hence increased ionic strength should increase the dimerization constant K. Figure 6 shows an autoradiogram for a gel in which [RNA] is held constant at 0.243 μM and ionic strength varies from 91.3 to 369 mM. The intensity profiles for low ionic strength show an intense double peak somewhat above the center of the scan, one or two much weaker peaks near the top, and as many as three narrow faster-running peaks near the bottom. The direction of electrophoresis being downward, the peaks near the top are the slowest-running species and, as discussed above, are assigned to the dimer. The very intense peaks are assigned to the monomer, and the peaks near the bottom must arise from fragmentation of the RNA.

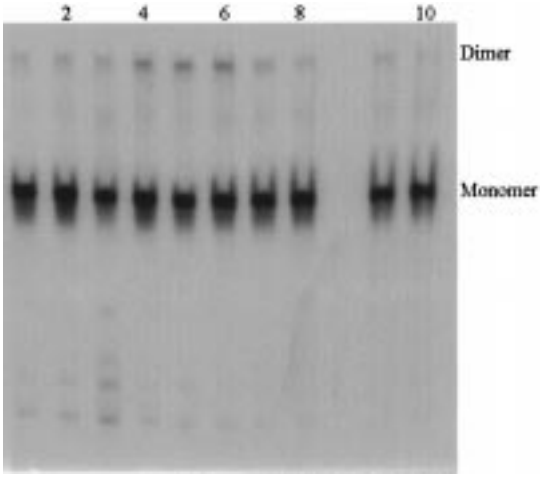

FIGURE 6: Autoradiogram of the 41-mer at 5 °C. RNA concentration is 0.243 μM throughout, but ionic strength is varied. Lanes 1−10 correspond to ionic strengths (in gel wells) of 91, 94, 102, 111, 119, 157, 192, 230, 300, and 369 mM, respectively.

Table 3: Spot Intensities for the 41-mer[a] as a Function of Ionic Strength I

I (mM)	run 1			run 2			run 3		
	slow[b]	center[c]	fast[d]	slow[b]	center[c]	fast[d]	slow[b]	center[c]	fast[d]
91.3	3.92	80.46	3.05	0.466	35.26	2.97	1.11	30.64	10.46
94.1	7.74	101.99	3.49	0.94	38.61	2.47	2.07	30.60	5.47
102.4	5.02	68.70	8.91	2.45	37.97	1.98	2.93	26.12	6.91
110.8	12.64	90.07	2.38	2.58	39.06	2.02	6.04	26.98	6.86
119.1	13.54	76.45	1.40	2.50	39.50	1.72	3.31	24.48	5.36
157.3	11.95	65.85	0.62	5.01	40.88	1.66	7.29	26.27	3.30
192.0	9.96	77.86	0.68	3.05	33.77	1.46	6.67	24.79	3.15
230.2	8.52	62.09	0.27	1.62	32.79	1.25	2.89	22.52	3.59
299.7	5.58	61.04	0.36	1.33	33.33	1.59	2.55	22.70	2.46
369.1	2.26	49.37	0.27	0.51	29.65	0.22	1.79	22.73	2.38

[a] Temperature = 5 °C. [b] Slowest-running peaks, near top of gel, arise from dimer. [c] Intense peaks at the center of the gel arise from monomer [d] Fastest-running peaks, near bottom of gel, arise from fragments of monomer

Table 4: Q[a] as a Function of Ionic Strength I for the 41-mer[b]

I (mM)	run 1	run 2	run 3	run 4	I (mM)	run 5[c]
91.3	0.0265	0.0073	0.0251	0.0620	86.8	0.0081
94.1	0.0422	0.0132	0.0421	0.0736	88.7	0.0588
102.4	0.0439	0.0360	0.0772	0.0811	94.2	0.0654
110.8	0.0819	0.0370	0.1655	0.1203	99.8	0.0398
119.1	0.1059	0.0351	0.0917	0.1049	105.4	0.0837
157.3	0.1081	0.0987	0.1946	0.0939	131.3	0.0317
192.0	0.0728	0.0512	0.1878	0.0722	153.5	0.0434
230.2	0.0784	0.0268	0.0828	0.0682	179.4	0.0232
299.7	0.0501	0.0217	0.0686	0.0688	227.6	0.0184
369.1	0.0241	0.0088	0.0466	0.0618	272.0	0.0163

[a] $Q = \theta_D\theta_T/(2\theta_M^2)$. [b] Temperature = 5 °C, [RNA] = 0.243 μM. [c] [RNA] = 0.261 μM for this run.

Table 3 gives measured spot intensities for three runs with the same RNA concentration and ionic strengths, differing in the conditions under which samples were equilibrated before loading into the gel well. As noted above, this pretreatment has no net effect because equilibrium is established in the well at the gel temperature, about 5 °C.

Table 4 gives calculated values of Q for the three gels of Table 3 and two others. The RNA concentration for the last

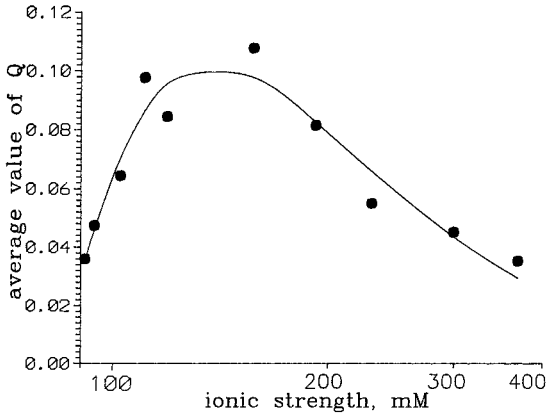

FIGURE 7: Dependence of Q on ionic strength (RNA concentration fixed) for the 41-mer at 5 °C (average value of five determinations). Solid curve is the best-fit four-parameter log-normal function.

run is 0.261 μM (instead of 0.243 μM for the other four) and the ionic strengths are also somewhat different, as shown. The values of K for ionic strengths near 113 and 183 mM are estimated by dividing the average values of Q for these ionic strengths by the RNA concentration plus 0.1 μM (L is probably somewhat smaller than 0.1). The average values for K are approximately 0.29×10^6 M^{-1} and 0.35×10^6 M^{-1}, respectively. The former is much closer to what was obtained from the results of Table 2 than to what was obtained from the results of Table 1.

Raising ionic strength (or raising salt concentration) should increase the amount of dimer relative to monomer. As ionic strength increases from 90 to 157 mM, the dimer peaks (near the top of the gel in Figure 6) do indeed grow in intensity relative to the monomer peaks (near the center); the cleavage peaks (at the bottom) diminish in intensity. For higher ionic strengths, however, the dimer peaks *decrease* in intensity relative to the monomer peaks, so that Q goes through a maximum. This is true for each of the five runs of Table 4. The average values of Q are shown in Figure 7, with a least-squares fit to a four-parameter log-normal function. The maximum is at $I = 137$ mM.

The ratio of intensities for the two slowest-running bands, equal to K' (equation 5), was also calculated as a function of ionic strength. The results are very imprecise, because both peaks are very small and difficult to measure accurately. However, the consensus result is the K' is 0.6 ± 0.3 for $I =$ 91 mM and decreases to 0.4 ± 0.2 for $I = 369$ mM. The total intensity in a given lane on the gel is constant for I from 91 to 157 mM but decreases for $I > 157$ mM.

For the 27-mer, several gels were prepared, run at 5 °C, and measured. They differ in the temperature at which samples were equilibrated before being loaded into the well of the gel. As is evident from Figure 8, this RNA produces a spot pattern similar to the 41-mer but the streak near the dimer is missing. This oligomer being considerably smaller than the 41-mer, it was expected that the rate constants for dimerization and dimer dissociation would be larger, allowing equilibrium to be established more quickly in solution. On the other hand, the existence of discrete monomer and dimer spots shows that interconversion between the forms in the gel is slow. This suggests that the streak near the dimer for the 41-mer is a second form of the dimer, not the result of interconversion between forms.

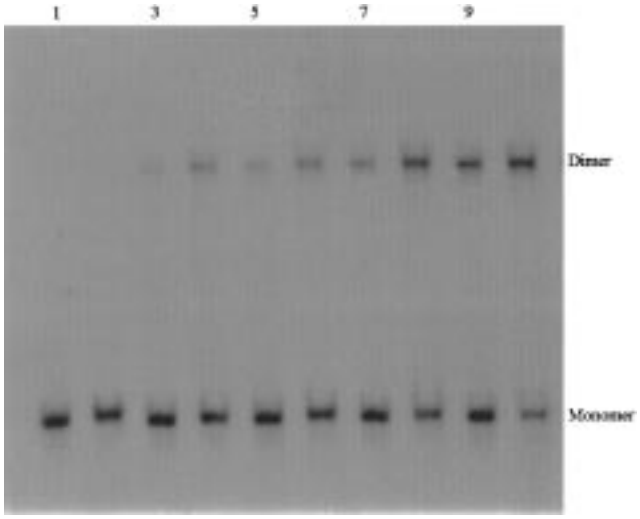

FIGURE 8: Autoradiogram of the 27-mer at 5 °C. Even and odd lanes have $I = 101$ and 170 mM, respectively. The lane number and [RNA] (in micromolar, except for radiolabeled) are as follows: lanes 1 and 2, 0 (only radiolabeled RNA); lanes 3 and 4, 0.085; lanes 5 and 6, 0.170; lanes 7 and 8, 0.513; lanes 9 and 10, 1.05.

Table 5: Spot Intensities and Q for 27-mer[a]

[RNA] (μM)	intensities, $I = 101$ mM				intensities, $I = 170$ mM			
	monomer	dimer	total[b]	Q	monomer	dimer	total[b]	Q
0	20.72	0	27.63	0	25.45	0	25.45	0
0.087	22.87	1.29	26.23	0.016	14.51	5.05	23.63	0.142
0.174	17.58	2.67	28.90	0.062	13.46	6.61	23.00	0.210
0.347	18.04	3.75	25.39	0.073	8.54	7.26	18.11	0.451
0.521	16.34	7.10	27.28	0.181	6.50	7.63	16.29	0.734
0.694	13.55	6.92	23.37	0.220	6.98	10.16	19.90	1.037
0.845	11.57	8.90	23.42	0.389	6.30	11.40	19.97	1.432

[a] Temperature = 5 °C, $Q = \theta_D\theta_T/(2\theta_M^2)$. [b] Total intensity includes peaks other than monomer and dimer.

Table 6: Spot Intensities and Calculated Q for 27-mer

[RNA] (μM)	intensities, $I = 101$ mM				intensities, $I = 170$ mM			
	monomer	dimer	total[b]	Q	monomer	dimer	total[b]	Q
0	12.00	0	12.00	0	8.72	0	8.72	0
0.087	13.14	0.62	13.76	0.025	7.74	2.58	10.33	0.223
0.181	15.79	1.14	16.93	0.039	6.95	2.60	9.55	0.257
0.524	12.82	2.61	15.42	0.122	4.04	4.83	8.86	1.314
1.038	7.50	3.58	11.08	0.353	3.66	8.26	11.92	3.674
1.354	6.91	4.73	11.64	0.578	2.88	11.42	14.30	9.820

Spot intensities and calculated values of Q for the 27-mer are given in Tables 5 and 6. Data were obtained for several RNA concentrations and for ionic strengths of 101 and 170 mM. The gel that gave the data in Table 5 showed some bands in addition to the monomer and dimer bands. Since the additional bands ran faster than the monomer or dimer bands, they were ascribed to cleavage products of the RNA. Their intensities were included in the total intensity (see Table 5) and therefore in the calculation of K. The assumption is that the presence of species other than monomer and dimer does not affect the monomer−dimer equilibrium.

The values of Q in Table 5, plotted vs nominal RNA concentration C of the 27-mer, give a slope, equal to K, of $(0.418 \pm 0.054) \times 10^6$ M^{-1} for $I = 101$ mM and $(1.628 \pm 0.094) \times 10^6$ M^{-1} for $I = 170$ mM. The values of r^2 for the linear fits are 0.984 and 0.923, respectively, and the values obtained for L, the concentration of RNA in the labeled

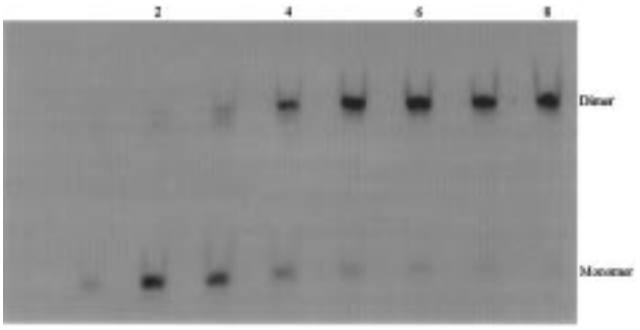

FIGURE 9: Autoradiogram of the 19-mer at 5 °C. Lanes 1−8 correspond to unlabeled RNA concentrations (in micromolar): 0, 0, 0.015, 0.087, 0.289, 0.521, 0.868, and 1.30. Lane 1 has 0.2 times the radiolabeled RNA of the other lanes and was not used in the analysis. Ionic strength is 118 mM.

Table 7: Results for 19-mer, $I = 118$ mM, 5 °C

[RNA] (μM)	monomer int	dimer int	total int	calcd Q
0	22.76	1.53	24.30	0.036
0.014	15.29	4.46	19.75	0.188
0.087	7.85	14.97	22.82	2.769
0.288	2.17	28.65	32.23	102.5
0.521	2.61	35.61	43.23	129.3
0.868	1.72	29.58	52.58	452.5
1.302	2.13	31.43	57.06	344.1

Table 8: Spot Intensities and Q, 19-mer, 20 °C

[RNA]	$I = 103$ mM			$I = 118$ mM			$I = 135$ mM		
(μM)	θ_M[a]	θ_D[b]	Q	θ_M[a]	θ_D[b]	Q	θ_M[a]	θ_D[b]	Q
0	6.160	0.029	0.0017	6.832	0.221	0.017	7.621	0.714	0.05
0.029	10.589	0.059	0.0028	7.381	1.444	0.117	5.645	3.205	0.45
0.058	6.690	0.115	0.0088	6.864	4.566	0.554	3.874	4.905	1.43
0.116	14.374	0.267	0.0095	5.713	6.883	1.328	2.492	8.741	7.91
0.174	17.314	0.341	0.0100	4.351	7.563	2.380	2.086	10.495	15.18
0.231	15.45	0.369	0.0122	3.824	8.115	3.313	1.364	5.099	8.86
0.289	14.103	0.500	0.0184	2.190	7.449	7.485	1.791	14.993	39.21

[a] Intensity of monomer. [b] Intensity of dimer.

oligomer, are zero within the error of the fit: (-0.059 ± 0.062) μM and (-0.030 ± 0.028) μM. For the data of Table 6, plots of measured Q vs. C turned sharply upward at the highest values of C, which we ascribed to the low, and hence difficult to measure, spot intensities for the monomer. Therefore, only the first 5 data points were used in the linear fit. The results for $I = 101$ mM were $K = (0.334 \pm 0.034) \times 10^6$ M^{-1} and $L = (-0.045 \pm 0.050)$ μM with $r^2 = 0.973$; for $I = 170$ mM, $K = (3.57 \pm 0.35) \times 10^6$ M^{-1} and $L = (-0.058 \pm 0.50)$ μM with $r^2 = 0.974$. In both cases L is essentially zero.

A third gel (data not shown) yielded, for 101 mM ionic strength, $K = (1.74 \pm 0.035) \times 10^6$ M^{-1} and $L = (-0.0057 \pm 0.0075)$ μM ($r^2 = 0.998$) and, for 170 mM ionic srength, $K = (6.11 \pm 0.59) \times 10^6$ M^{-1} and $L = (-0.037 \pm 0.045)$ μM ($r^2 = 0.955$). A fourth gel (data not shown) yielded $K = (0.234 \pm 0.037) \times 10^6$ M^{-1} with $L = (-0.034 \pm 0.102)$ μM for 101 mM ionic strength ($r^2 = 0.909$) and $K = (3.02 \pm 0.48) \times 10^6$ M^{-1} with $L = (-0.078 \pm 0.102)$ μM for 170 mM ionic strength ($r^2 = 0.907$). In all cases L is essentially zero.

An autoradiogram of a gel for the 19-mer is shown in Figure 9, and results for this oligomer are given in Tables 7 and 8. Other experiments (results not shown here) had previously shown that intensity patterns did not depend on incubation temperature of the samples before loading or on electrophoresis time. The latter (load-and-run) experiments showed that reequilibration (dimerization and dissociation of dimer) did not occur to an appreciable extent in the gel.

The gel giving the data of Table 7 was run at 5 °C; the samples contained 100 mM NaCl and 1× TBE buffer, so that the total ionic strength in the well was 118 mM. To have large enough intensities for accurate measurement for both monomer and dimer spots, measurements were made on two different exposures of the autoradiogram. Still, the measured monomer spot intensities for 0.87 and 1.302 μM were too small to be accurately measured. Using the five lowest concentration points for a linear fit of Q vs known RNA concentration gave a low value of r^2 (0.87) and $K = (277 \pm 44) \times 10^6$ M^{-1}. The value of L was very close to zero. A second gel with the same RNA concentrations and ionic strength was run (data not shown). Only the five lowest concentration points gave meaningful monomer intensities. Using these in the linear fit gave ($r^2 = 0.95$) $K = (239 \pm 40) \times 10^6$ M^{-1}. The value of L was essentially zero.

The results in Table 8 were obtained from gels run at 20 °C. Samples contained 1×TBE buffer and had [NaCl] = 0, 100, and 200 mM, which give ionic strengths in the gel well of 103, 118, and 135 mM, respectively. It was expected that the increased temperature would lower the dimerization constant, and that the dimerization constant would increase with ionic strength. By using intensities from two exposures of the autoradiogram, we were able to reliably measure both dimer and monomer spot intensities.

A linear fit of Q vs C for ionic strength 103 mM gives ($r^2 = 0.88$) $K = (0.050 \pm 0.008) \times 10^6$ M^{-1} and $L = (0.054 \pm 0.042)$ μM. For ionic strength 118 mM, the highest concentration point was dropped (monomer intensity was inaccurately measured) before the same treatment was performed. The result was ($r^2 = 0.99$) $K = (14.8 \pm 0.8) \times 10^6$ M^{-1} and $L = (-0.014 \pm 0.011)$ μM. For ionic strength 135 mM, the two highest RNA concentration points were not used in the linear fit, which resulted in ($r^2 = 0.94$) $K = (137 \pm 18) \times 10^6$ M^{-1} and $L = (-0.032 \pm 0.031)$ μM. The reason for dropping high-concentration points was that the low monomer intensity, subject to large measurement errors, made Q very uncertain.

DISCUSSION

With all the oligonucleotides considered, the temperature at which the RNA was held before before loaded into the gel had no effect on the intensity patterns seen after electrophoresis. Other measurements, some of which are discussed here, show that these patterns do depend on the running temperature of the electrophoresis experiment. Therefore, the interconversion and diffusion rates in bulk solution at 5 °C are fast enough so that the monomer−dimer equilibrium is established during the 10−15 min required for RNA to move out of the well and into the gel proper.

However, the interconversion between forms in the gel at 5 °C is slow relative to the migration rate. The proof of this is that distinct bands for the monomer and dimer are observed. When electrophoresis was carried out at 30 °C, interconversion rates in the gel were so high that only a single band was seen for all sequences studied (data not shown). Simulations show that a single band arises when intercon-

Table 9: Determined Values of Dimerization Constants

oligomer	ionic strength *I* (mM)	temperature (°C)	constant *K* (in 10^6 M^{-1})
41-mer	113	5	0.069
41-mer	183	5	0.105
41-mer	105	5	0.147
41-mer	113	5	0.29
41-mer	183	5	0.35
27-mer	101	5	0.42
27-mer	170	5	1.63
27-mer	101	5	0.33
27-mer	170	5	3.57
27-mer	101	5	1.74
27-mer	170	5	6.11
27-mer	101	5	0.23
27-mer	170	5	3.02
19-mer	118	5	277
19-mer	118	5	239
19-mer	100	20	0.050
19-mer	118	20	14.8
19-mer	135	20	137

version between forms is faster than the migration rate in the gel (*17*). Interconversion rates are expected to be slower for longer RNA segments, which may allow measurements of equilibrium constants at higher temperatures for such segments.

Nondenaturing gel electrophoresis shows that the three sequences studied exist in two major forms, the relative amounts of which vary with concentration and ionic strength. Usually, the fastest running band in the gel is the monomer (in the stem−loop form) while the much slower running band or bands is the dimer. Sometimes, very fast running bands with low intensity appear; these are ascribed to fragmentation of the monomer RNA arising from handling and the age of the sample.

In the case of the 41-mer the monomer can exist in more than one form, as revealed in the autoradiogram by the presence of several overlapping bands that are part of the monomer envelope. Since care was taken to isolate and label oligomers, we believe the complex shape of the monomer band reflects differences in secondary structure rather than heterogeneity in length or sequence of the RNA. The presence of more than one monomer band was earlier observed for a stem-loop DNA 12-mer studied by nondenaturing gel electrophoresis (*17*).

From electrophoresis measurements alone it is not possible to tell if the dimer is in the kissing or duplex form or if the band is due to a rapidly interconverting mixture of the two. For the 41-mer, the dimer band appears to consist of two bands, suggesting that both kissing and duplex forms are present. Interconversion between the two dimer forms should be slowest for this oligomer.

Table 9 summarizes the determined values of the dimerization equilibrium constant. The constant *increases* by several orders of magnitude as the length of the stem of the monomeric RNA *decreases* by a factor of 2, from the 41-mer to the 19-mer. Thus ΔG for dimerization decreases markedly with length. It is unlikely that ΔH changes much for the following reasons: When the kissing dimer is formed, the stems of the monomers are unchanged; when the duplex dimer is formed, the stems melt to form an equivalent structure in the dimer. In either case, bases in the stem should contribute little to ΔH. Since all three RNAs have the same palindrome in the loop, base paring in either form of the dimer should give similar values of ΔH for all three, making the enthalpy difference between reactions $\Delta\Delta H \sim 0$. Thus a decrease in ΔS must be responsible for the increase in ΔG.

Consider ΔS for joining two rods of equal length. The partition function for a rotating or orientating rod (*27*) is $8\pi^2\ IkT/h^2$, where I is the moment of inertia, equal to $\rho L^3/12$, where ρ is the mass per unit length and L is the rod's length. The rotational or orientational entropy for one such rod is

$$k + k \ln q = k + k \ln\left[\frac{8\pi^2 kT\rho L^3}{12h^2}\right]$$

The molar entropy change for joining two rods of length L to form one rod of length $2L$ is therefore

$$\Delta S = R\left[1 + \ln\left[\frac{2\pi^2 kT\rho(2L)^3}{3h^2}\right]\right] - 2R\left[1 + \ln\left[\frac{2\pi^2 kT\rho L^3}{3h^2}\right]\right] = -R - R \ln\left[\frac{2\pi^2 kT\rho}{3h^2}\right] + R \ln\left[\frac{8L^3}{L^6}\right]$$

Comparing a rod of length rL with a rod of length L, the difference in molar entropies of dimerization is

$$\Delta\Delta S = R \ln[(\mathrm{rL})^{-3}] - R \ln [L^{-3}] = -3R \ln (r)$$

The orientational contribution to ΔH is just $-R$, independent of L, so that the contribution to $\Delta\Delta H$ is zero. Since the dimerization equilibrium constant is proportional to $e^{\Delta S/R}$, the ratio of dimerization constants for the two rods is $e^{\Delta\Delta S/R} = r^{-3}$. Doubling the rod length would decrease the dimerization constant by an order of magnitude.

Another reason for the decrease of ΔS with length is probably the increased flexibility of the monomeric RNA. This should lead to a large increase in the number of conformers (secondary structures), which is consistent with the complicated shape of the monomer band in the traces of Figure 4. The increase in flexibility, and hence in the number of conformers, should be less important for the dimer than for the monomer, so that the entropy of the monomer should increase with length more than the entropy of the dimer. Thus the conformational contribution to ΔS, like the orientational contribution, should decrease with length, giving an important negative contribution to $\Delta\Delta S$ and decreasing the dimerization constant.

Previous NMR studies have shown that the 19-mer used in this study and a 24-mer from the dimer initiation site of the LAI strain of the HIV virus adopt a stable kissing structure in solution (*14*, *15*). In our studies we are unable to confirm the kissing structure for the 19-mer and have no knowledge of the structure of the dimers formed for the 27- and 41-mers. However, in the case of the 41-mer, the electrophoresis gels show two broad bands in the dimer region, Figure 4, suggesting that this RNA may exist in two dimer forms under the conditions of the experiment. The two band areas were combined for the calculation of the dimerization constant.

The slower-moving band was generally sharper and more intense than the other. It is suggested that this band arises from the duplex or tight dimer, which should have a more

rigid structure than the kissing dimer. Furthermore, the kissing dimer may be able to rearrange in the gel either to the tight dimer or to monomers, which would explain why the faster-moving dimer band is broader than the other. The small areas of both peaks makes it difficult to obtain accurate values for the ratio, which is the equilibrium constant K' of eq 5. The values obtained for the ratio of kissing to duplex dimer were 0.48 ± 0.22 at ionic strength 113 mM and 0.91 ± 0.35 at ionic strength 413 mM.

The dependence of the dimerization constant K on ionic strength has been studied, mostly for the 41-mer. The value of this constant increases with ionic strength, as expected since the dimer is formed from two similarly charged monomers, but only up to an ionic strength of about 140 mM (Tables 3 and 4). Above this, K apparently decreases with ionic strength, so that its value at 400 mM is close to its value at 100 mM. However, the intensity of dimer and monomer bands is also observed to decrease when ionic strength exceeds 140 mM, suggesting that RNA is lost before it enters the gel. Since the quantity that is actually calculated from the data is KC, where C is the total RNA concentration, loss of RNA would give an apparent decrease in K. If higher ionic strength facilitates dimerization, it may also lead to formation of higher order oligomers that do not enter the gel, leading to the decrease in total intensity. We have occasionally found radioactivity in the sample tubes after transferring sample to the gel.

The nature of the dimer that forms from two monomeric RNAs in the DIS of HIV-1 has been extensively addressed in the literature (*2*, *7*, *8*). These studies often used long RNAs that could potentially adopt a variety of different forms, making it difficult to analyze the system in terms of a simple two-state model. In fact, gels for these RNAs often show the presence of multiple species. Although dimers have been found in these studies, and a kissing structure having the palindromic regions of two monomers hydrogen-bonded to each other is a reasonable intermediate, direct evidence for such a structure with large RNAs has not been forthcoming. Paillart et al. (*28*) measured the equilibrium dissociation constants of long RNAs (615-mers) from the DIS of HIV. They reported dissociation constants, K_d, of $\sim 20 \times 10^{-9}$ M (K, 5×10^7 M^{-1}) and suggested that the dimer is in the kissing form. The enthalpic contribution to the driving force for the formation of either form of the dimer, kissing or duplex, is base pairing of the nucleotides in SL1 of the monomer. This should be negative. The entropic contribution is more difficult to predict but contributions to orientational entropy and the presence of multiple forms of the monomer *decrease* the dimerization constant as RNA length increases. While the packaging region is significantly larger than the 41-mer, the observed trends in the dimerization constant suggest that K for the former should be smaller than $\sim 10^5$ M^{-1}.

An important driving force for dimer formation appears to be binding to the nucleocapsid protein, NCp7. The protein is known to stabilize dimeric retroviral RNA (*6*, *29*, *30*) and it is known to convert the kissing dimer of HIV-1 LAI into the duplex dimer (*13*). We are in the process of measuring the monomer−dimer equilibrium constants of large segments of HIV RNA in the presence and absence of nucleocapsid protein and drugs to determine the extent to which the packaging region may be used as a target for new treatment strategies for AIDS. The results of these studies will be reported in due course.

SUMMARY AND CONCLUSIONS

We have measured dimerization constants for sequences of HIV-1 RNA using nondenaturing polyacrylamide gel electrophoresis. A standard DNA sequencing apparatus is used to separate RNA forms, and their amounts are measured by scanning autoradiographic spot intensities. If the interconversion rate between forms in the gel is slow relative to the migration rate, peak areas are proportional to the amounts of forms that enter the gel from the loading well. However, for systems of the type studied here that can interconvert rapidly in solution, introduction of the equilibrated sample into the well dilutes the sample and alters the equilibrium to one determined by the conditions in the well. Concentrations and ionic strengths in the well must then be estimated, leading to lower than expected reproducibility in the determined values of K.

It is shown that an increase in the length of the sequence of RNA from 19 to 41 leads to a decrease in the dimerization constant by about 3 orders of magnitude. The difference in dimerization free energy with length is due in part to changes in rotational entropy for the rodlike molecules and the existence of multiple forms of the monomer as RNA length increases. The observation that the dimerization constant decreases with increased length strongly suggests that K for the full-length packaging signal is small and that factors other than sequence (such as the presence of protein) must contribute to stabilizing the duplex form of the HIV genome in the virion.

ACKNOWLEDGMENT

Helpful discussions with, and advice from, Professor P. N. Borer are gratefully acknowledged.

REFERENCES

1. Coffin, J. (1985) in *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H., and Coffin, J., Eds.) Vol. 2, pp 17−74, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Laughrea, M., and Jetté, L. (1994) *Biochemistry 33*, 13464−13474.
3. Clever, J., Sassetti, C., and Parslow, T. G. (1995) *J. Virol. 69*, 2101−2109.
4. Clever, J. L., and Parslow, T. G. (1997) *J. Virol. 71*, 3407−3414.
5. Harrison, G. P., and Lever, A. M. (1992) *J. Virol. 66*, 4144−4153.
6. Darlix, J. L., Gabus, C., Nugeyre, M. T., Calvel, F., and Barré-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−699.
7. Paillart, J.-C., Marquet, R., Skripkin, E., Ehresmann, B., and Ehresmann, C. (1994) *J. Biol Chem.* 269, 27486−27493.
8. Muriaux, D., Girard, P.-M., Bonnet-Mathoniere, B., and Paoletti, J. (1995) *J. Biol. Chem. 270*, 8209−8216.
9. Pappalardo, L., Kerwood, D. J., Liu, C., Pelczer, I., and Borer, P. N. (1998) *J. Mol. Biol. 288*, 801−818.
10. McBride, M. S., and Panganiban, A. T. (1996) *J. Virol. 70*, 2963−73.
11. Laughrea, M., and Jetté, L. (1996) *Biochemistry 35*, 1589−1598.
12. Laughrea, M., and Jetté, L. (1997) *Biochemistry 36*, 9501−9508.
13. Mariaux, D., De Rocquigny, H., Roques, B. P., and Paoletti, J. (1996) *J. Biol. Chem. 271*, 33686−33692.
14. Mujeeb, A., Clever, J. L., Billeci, T. M., James, T. L., and Parslow, T. G. (1998*) Nat. Struct. Biol. 5*, 432−6.

15. Dardel, F., Marquet, R., Ehresmann, C., Ehresmann, B., and Blanquet, S. (1998) *Nucleic Acids Res. 26*, 3567−3571.
16. Skripkin, E., Pillart, J. C., Marquet, R., Blumenfeld, M., Ehresmann, B., and Ehresmann, C. (1996) *J. Biol. Chem. 271*, 28812−28817.
17. Shubsda, M., Goodisman, J., and Dabrowiak, J. C. (1999) *Biophys. Chem 76*, 95−115.
18. Grodberg, J., and Dunn, J. J. (1988) *J. Bacteriol. 170*, 1245.
19. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51.
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
21. Wyatt, J. R., Chastain, M., and Puglisi, J. D. (1991) *BioTechniques 11*, 464−469.
22. Borer, P. N. (1972) Ph.D. Dissertation, University of California, Berkeley.
23. Sugiyama, T., and Fraenkel-Conrat, H. (1961) *Proc. Natl. Acad. Sci. U.S.A. 47*, 1393−1397.
24. Cann, J. R. (1996) *Electrophoresis 237*. 1−16.
25. Elsinger, J., and Blumberg, W. E. (1973) *Biochemistry 12*, 3648−3660.
26. Fried, M., and Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505−25.
27. Goodisman, J. (1997) *Statistical Mechanics for Chemists*, pp 84−5, John Wiley and Sons, New York.
28. Pailart, J.-C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5572−5577.
29. DeRocquigny, H., Gabus, C., Vincent, A., Fournie-Zaluski, M.-C., Rogues, B., and Darlix, J.-L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6472−6476.
30. Sakagushi, K., Zambrano, N., Baldwin, E. T., Shapiro, B. A., Erickson, J. E. Omichinski, J. G., Clore, G. M., Gronenborn, A. M., and Appela, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5219−5223.

BI990744T