Pergamon

Bioorganic & Medicinal Chemistry 10 (2002) 3663–3672

BIOORGANIC & MEDICINAL CHEMISTRY

Footprinting and Circular Dichroism Studies on Paromomycin Binding to the Packaging Region of Human Immunodeficiency Virus Type-1

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

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

Received 15 February 2002; accepted 12 April 2002

Abstract—We have studied the interaction of the aminoglycoside drug, paromomycin, with a 171-mer from the packaging region of HIV-1 (ψ-RNA), using quantitative footprinting and circular dichroism spectroscopy. The footprinting autoradiographic data were obtained by cutting end-labeled RNA with RNase I or RNase T1 in the presence of varying paromomycin concentrations. Scanning the autoradiograms produced footprinting plots showing cleavage intensities for specific sites on the ψ-RNA as functions of drug concentration. Footprinting plots showing binding were analyzed using a two-state model to give apparent binding constants for specific sites of the ψ-RNA. These plots show that the highest-affinity paromomycin binding site involves nucleotides near bulges in the main stem and SL-1, and other nucleotides in SL-4 of the ψ-RNA. RNase I gives an apparent value of K for this drug site of $\sim 1.7\times10^5$ M^{-1} while RNase T1 reports a value of K of $\sim 8\times10^4$ M^{-1} (10 mM Tris HCl, pH 7). Footprinting shows that loading the highest affinity site with paromomycin causes structural changes in the single-stranded linker regions, between the stem-loops and main stem and the loops of SL-1 and SL-3. Drug-induced structural changes also affect the intensity of the 208 nm band in the circular dichroism spectrum of the ψ-RNA. Fitting the changes in CD band intensity to a two-state model yielded a binding constant for the highest-affinity drug site of 6×10^6 M^{-1}. Thus, the binding constants from footprinting are lower than those obtained for the highest-affinity site from the circular dichroism spectrum, and lower than those earlier obtained using absorption spectroscopy (Sullivan, J. M.; Goodisman, J.; Dabrowiak, J. C., *Bioorg. Med. Chem. Lett*. 2002, *12*, 615). The discrepancy may be due to competitive binding between drug and cleavage agent in the footprinting experiments, but other explanations are discussed. In addition to revealing sites of binding and regions of drug-induced structural change, footprinting showed that the loop regions of SL-1, SL-3 and SL-4 are exposed in the RNA, whereas the linker region between SL-1 and SL-2 is 'buried' and not accessible to cutting by RNase I or RNase T1.

Introduction

The aminoglycosides are drugs that exhibit antibacterial effects by binding to the A site of ribosomal RNA.[1,2] Puglisi and co-workers[3–5] used NMR to show that the aminoglycoside paromomycin (Fig. 1) binds to a bulged region of A-site RNA. Ring I of the drug occupies a binding pocket created by a non-canonical A·A base pair, while the amino groups on the 2-deoxysteptamine ring (ring II) of paromomycin are involved in hydrogen bonds with heterocyclic bases that form the binding pocket. A crystal structure of paromomycin bound to a novel RNA construct containing two A-sites has also recently been published. The structure shows paromomycin bound in an enlarged deep groove created by two bulging and one unpaired adenines with the drug extensively hydrogen bonded through water to RNA.[6] Although NMR[7–10] and various other techniques[11–18] have been used to study aminoglycoside–RNA interactions, a clear picture of the factors that control binding specificity has not yet emerged. In general, the drugs prefer to bind to bulged or other non-Watson–Crick secondary elements of RNA. This is mainly because the drugs are too large to fit into the grooves of regular A-form RNA structures. The aminoglycosides have a number of protonated amine groups, the spatial orientation of which can be easily changed. Since these groups are often involved in hydrogen bond and/or electrostatic interactions with sites on RNA, the drugs

*Corresponding author. Tel.: +1-315-443-4601; fax: +1-315-443-4070; e-mail: jcdabrow@syr.edu

0968-0896/02/$ - see front matter
PII: S0968-0896(02)00220-1

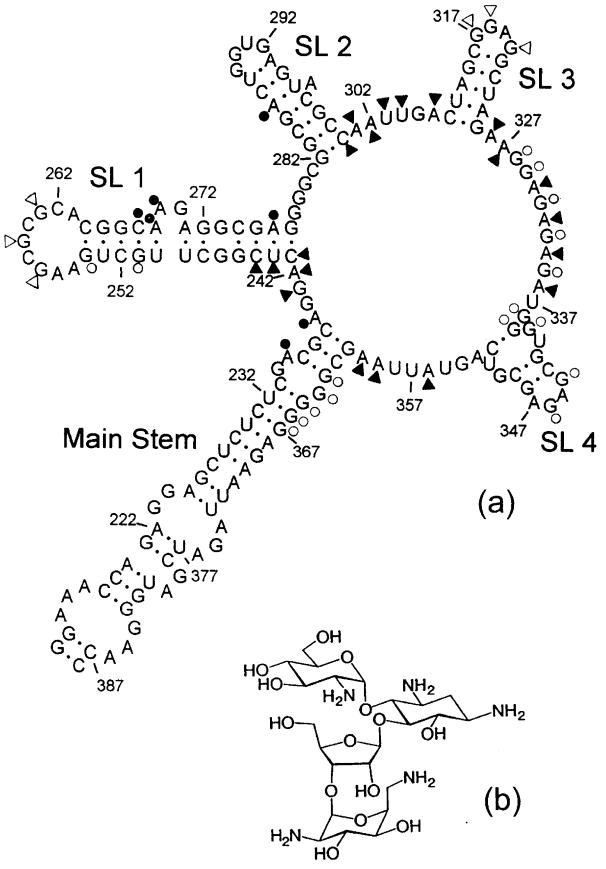

Figure 1. (a) Model of Ψ-RNA from the packaging region of HIV-1 (LAI) with a summary of binding/enhancement information for paromomycin. The RNA consists of the 'main stem', positions 213–238 and 361–388; SL-1, which contains the Dimer Initiation Site; SL-2, having the 5′ splice donor site; SL-3, and SL-4, the latter containing the start codon (AUG) for the *gag* gene. Nucleotides involved in binding and enhancement (structural changes) for RNase I are shown as filled circles and triangles, respectively. Nucleotides involved in binding and enhancement (structural changes) for RNase T1 are shown as open circles and triangles, respectively. (b) Structure of paromomycin.

can adapt to a variety of different binding pockets on RNA.

In an effort to explore the structural features of the drug that affect the affinity and specificity, Tor and co-workers chemically modified a number of aminoglycosides by attaching intercalators and reporting groups[19,20] and by forming drug dimers.[21] These analogues selectively inhibit the catalytic effects of ribozymes, and they are able to prevent proteins from binding to their interaction sequences on RNA. Tor, Goodman and co-workers[22] also synthesized the guanidinum analogues of tobramycin and neomycin B. These analogues were found to be about 100 times more effective than the parent aminoglycosides in inhibiting replication of the HIV virus in HeLa cells.

Although RNA is an attractive target for drugs,[23] relatively little is known about the way in which it can bind small ligands. Potential RNA targets are often large molecules that can fold into complex structures having many binding pockets for drugs. Measuring the affinities of drugs for various sites and identifying the structural features of RNA that form sites is critically important for designing agents that can selectively bind to one RNA in the presence of a number of competing RNAs. In an earlier report,[24] we outlined the theory and experimental protocols needed to obtain binding constants from drug–RNA footprinting data, using as an example the aminoglycoside drug paromomycin, and ψ-RNA from the packaging region of HIV-1 (LAI), (Fig. 1). Since this segment of RNA controls a number of important steps in the life cycle of the virus,[25] modifying its function with drugs may be an effective way of blocking the spread of the virus that causes AIDS. We have found that the absorption spectrum of ψ-RNA is sensitive to the presence of paromomycin and other aminoglycoside drugs.[26] By fitting the absorbance changes at 260 nm to a two-state model and assuming that, at low drug concentrations, a single drug binds to the highest affinity binding site, the binding constant was found to be $1.8\times10^{6}\ M^{-1}$.

In this report, we quantitatively analyze RNase I and RNase T1 footprinting data for paromomycin interacting with ψ-RNA obtaining apparent binding constants from the spot intensities appearing on the footprinting autoradiogram. The footprinting data are compared with the paromomycin-induced changes in the circular dichroism spectrum of ψ-RNA.

Results and Analysis

Theory

In the quantitative footprinting experiment involving drug binding to RNA[24] or DNA[27,28] one obtains footprinting plots showing how the intensity of a given band on the autoradiogram, corresponding to the rate of cleavage at a particular nucleotide position, changes with drug concentration. The rate of cleavage, R_i, is equal to the product of the concentration of the nucleotide, $[S_i]$, the concentration of the cutting agent at the nucleotide position, $[C_i]$, and the rate constant associated with cleavage (in this case, hydrolysis) at the site, k_i.

$$R_i = k_i[S_i][C_i] \tag{1}$$

For RNase I, the rate constant k_i, is relatively large for single-stranded RNA, particularly at A, while, for RNase T1, k_i is large for sites with unpaired G.[29]

The concentration of a particular nucleotide position that is not covered by drug, $[S_i]$, is related to the total concentration of nucleotides at that position, $[S_i]_T$, by the drug binding equilibrium:

$$K_i = \frac{[S_i]_T - [S_i]}{[S_i][D]} \tag{2}$$

Here, K_i is the equilibrium constant for paromomycin binding at a specific nucleotide i, [D] is the concentration

of free drug, and $[S_i]_T-[S_i]$ is the concentration of sites i with drug bound. Since there is one site i on each oligomer, $[S_i]_T$ is equal to the oligomer concentration S_T. The free-drug concentration is the total drug concentration minus the concentration of drug bound at all sites on the oligomer. Since $[C]_i$ is equal to the total concentration of cleavage agent minus the concentration of cleavage agent actually involved in the cleavage process, $[C]_i$ is expected to be the same for all sites. Writing it as [C], we have

$$R_i = k_i[C]\frac{S_T}{1 + K_i[D]} \quad (3)$$

The intensity of the spot on the gel corresponding to cleavage at site i, I_i, is proportional to R_i. Eq. 3 explains footprinting plots associated with sites at which binding is taking place and spot intensities decrease with added drug, called type 1 plots. If the free-drug concentration were known, one could obtain site-specific binding constants K_i by fitting plots of spot intensity versus free-drug concentration to the Langmuir form, $I_i = a_i(1 + K_i[D])^{-1}$. However, there are two potential problems with extracting binding constants from footprinting data.

One is that, in the footprinting experiment, usually only the total drug concentration, D_T, is known. Since [D] is a monotonic function of D_T, it is possible to use the same Langmuir functional form and fit the type 1 footprinting plots to

$$I_i = \frac{a_i}{1 + b_i D_T} \quad (4)$$

Larger values of b_i can be shown to correspond to larger values of K_i. If it is known that only a single binding event is taking place (as for drug concentrations for which only the primary site is loading), $[S_i]_T-[S_i]$ is equal to the concentration of bound drug, which is $D_T-[D]$. This two-state model makes eq 2 become

$$K_i = \frac{S_T - [S_i]}{[S_i](D_T - S_T + [S_i])} \quad (5)$$

which is a quadratic equation to be solved for $[S]_i$, given K_i, D_T and $S_T = [S_i]_T$. Then one can substitute $[S_i]$ for each D_T into eq 1 and find the values of K_i and $k_i[C]$ which give the best fit to the footprinting plot. For secondary sites, [D] is smaller than $D_T-S_T+S_i$, and the model underestimates K_i. Another problem with binding constants from footprinting data is related to the cutting agent. If it competes with the drug for binding to a site, the drug binding constant from footprinting will be too low. This is not a problem if the cutting agent is non-specific (binds and cuts at many sites), and if the ratio of polymer sites covered by drug to those that are free is small. This situation is found for many quantitative drug-DNA footprinting experiments involving the non-specific enzyme, DNase I, and relatively long fragments ($\sim$200 bp) of DNA.[27,28,30] However, if the cutting agent can bind at only a few sites, and its binding constant is comparable to that of the drug, the drug and the cutting agent effectively compete for the same site, and footprinting data will give a value of K that is *lower* than the true value.[31]

A drug-induced increase in cutting at a site, called an 'enhancement', can also occur in footprinting experiments. If all of the cutting sites are not covered with the cutting agent, drug binding to the polymer will displace the cutting agent to other sites, which increases R_i in eq 1 by increasing $[C_i]$. However, R_i can also be increased without displacement if drug causes a structural change in the polymer, changing k_i in eq 1. For single-stranded RNAs, which can readily adopt a variety of secondary structures, enhancements in drug-RNA footprinting experiments are most likely to occur by the latter mechanism.

Footprinting data

The conditions of the RNase I and RNase T1 footprinting experiments, given in an earlier report,[24] are briefly summarized in the Experimental of this work. The footprinting autoradiograms, Figures 2 and 3, were scanned and intensities of bands, each corresponding to cleavage of RNA at different sites on ψ-RNA, were measured. Footprinting plots were of several types: type 1, band (cutting) intensity decreases with drug concentration; type 2, intensity increases with drug concentration up to some maximum and then decreases; type 3, intensity increases with drug concentration; type 4, intensity hardly changes with drug concentration. When compared to RNase I, RNase T1 had fewer cutting sites on the ψ-RNA. Most of the plots for RNase T1 were type 1, but, in the type 2 plots, decreases occurred at higher drug concentration than for RNase 1. Selected footprinting plots from the RNase I and RNase T1 experiments are shown in Figures 4 and 5, respectively.

RNase I. Representative footprinting plots for RNase I cleaving ψ-RNA in the presence of paromomycin are shown in Figure 4. Apparent binding constants for paromomycin, derived from footprinting data, are given in Table 1. The binding constants obtained with RNase I fall into two groups that differ in the apparent value of K_i by about a factor of 2.

Type 1 plots were found for sites A235 and A239, and for site A276, giving values of K_i of 1.6×10^5 M^{-1} and 1.7×10^5 M^{-1}, respectively. (The values of K_i were determined to minimize R, the sum of the squared deviations between observed and calculated spot intensities on all three footprinting plots.) These nucleotides

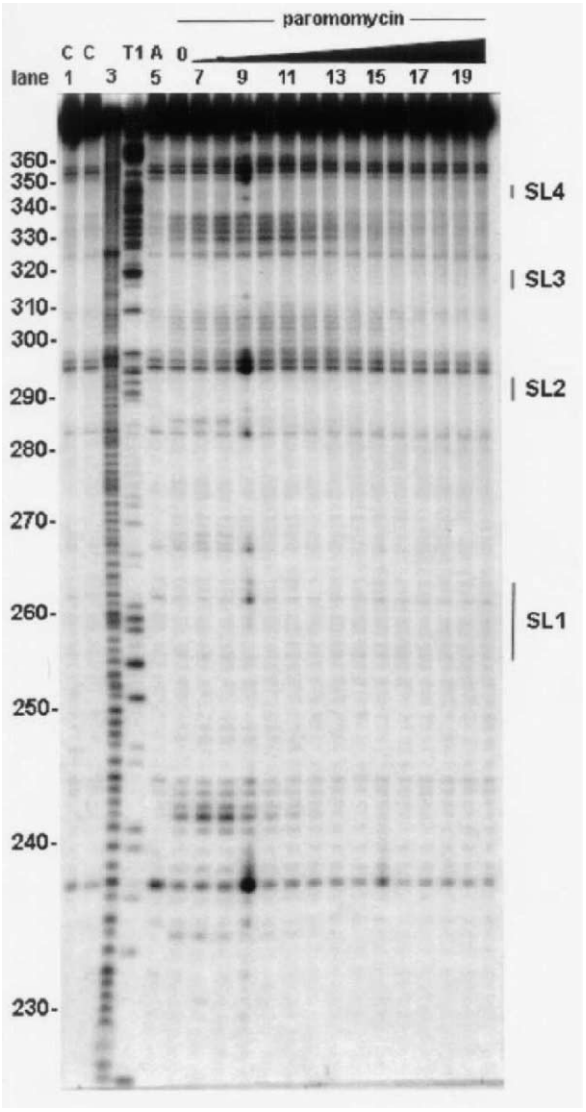

Figure 2. Footprinting autoradiogram of cleavage of Ψ-RNA using RNase I in the presence of paromomycin. Lanes 1 and 2, RNA alone; lane 3, hydroxide ion cleavage ladder; lane 4, RNase T1 ladder (G); lane 5, RNase A ladder (C, U). The lane number and concentration in micro molar of paromomycin respectively are: 6, 0; 7, 2.5; 8, 5; 9, 7.5; 10, 10; 11, 12.5; 12, 15; 13, 17.5; 14, 20; 15, 25; 16, 30; 17, 35; 18, 40; 19, 45; 20, 50. The sequence of the RNA is shown in Fig. 1. The footprinting experiments were carried out in the buffer, 10 mM Tris–HCl (pH 7).

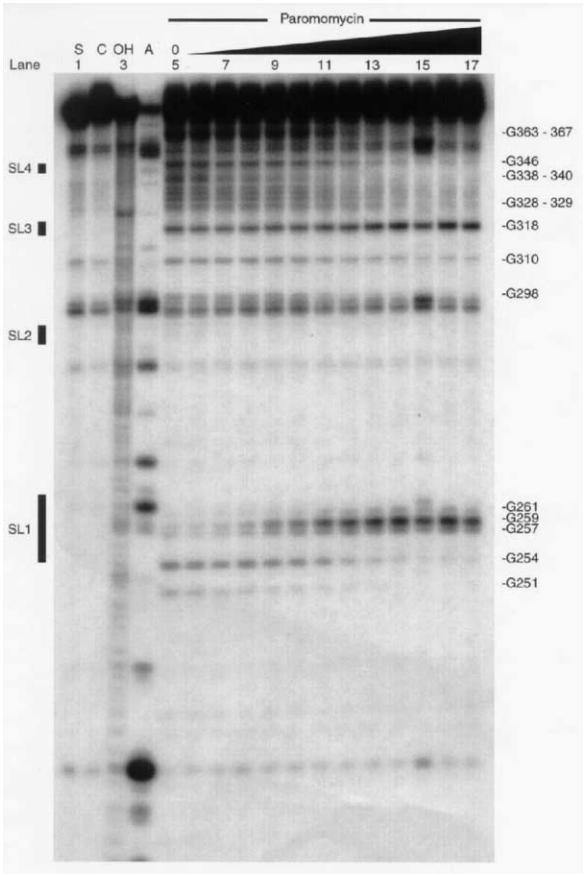

Figure 3. Footprinting autoradiogram of cleavage of Ψ-RNA using RNase T1 in the presence of paromomycin. Lanes 1 and 2, RNA alone; lane 3, cleavage of RNA by hydroxide; lane 4, RNase A cleavage ladder. The lane number and concentration of paromomycin (micromolar), respectively, are 5, 0; 6, 5; 7, 10; 8, 15; 9, 20; 10, 25; 11, 30; 12, 35; 13, 40; 14, 45; 15, 50 (not scanned); 16, 55; 17, 60.

are located at or near bulges in SL-1 and the main stem. This enzyme also detects drug binding at C267, A268, and A269, in SL-1, and at A286, in SL-2, with values of K_i of 8×10^4 and 6×10^4 M^{-1}, respectively. These nucleotides, at which cutting is weak, are located in regions of the ψ-RNA that are predicted to have duplex character (see Fig. 1).

The footprinting plots for 267, 268 and 269, which reflect the same binding event, were analyzed together using a single value of K_i (but a different value of the parameter k_i[C] for each—see eq 3). The uncertainties in the K_i were estimated at almost 50%, so they could be identical; both are significantly different from K_{235}. The error in K_{235} was estimated by fixing K_{235} at a value different from the best value and minimizing R, the sum of the squared deviations, with respect to the parameter k_{235}[C]. We found the change in K_i which increased the minimum R by 10% over the best value. The result was $K_{235} = (1.6 \pm 0.5)\times10^5$ M^{-1}.

Most of the footprinting plots observed for RNase I are type 2, for which cutting increases to a maximum followed by a decrease. These mainly occur in regions of the ψ-RNA that are predicted to be single-stranded, that is, 241–245 (between the main stem and SL-1), 300–311 (between SL-2 and SL-3), 326- 336 (between SL-3 and SL-4), and at A356, A359, A360 (between SL-4 and the main stem).

RNase T1. Typical footprinting plots for RNase T1 are shown in Fig. 5. The type 1 footprinting plots for RNase T1 return apparent values of K_i for paromomycin that are in general lower than those obtained with RNase I, Table 1. The drug site with highest affinity, K_i $\sim7\times10^4$ M^{-1}, is at positions G338, G339, G340, G344, and G346, in the start codon of the *gag* gene and in SL-4. This enzyme also detects a paromomycin binding site at positions G328, G329, G331, G333, and G335 with an apparent value of K_i $\sim7\times10^4$ M^{-1}. RNase T1 also finds two 'low K' paromomycin binding events at G251 and G254, in the stem of SL-1,

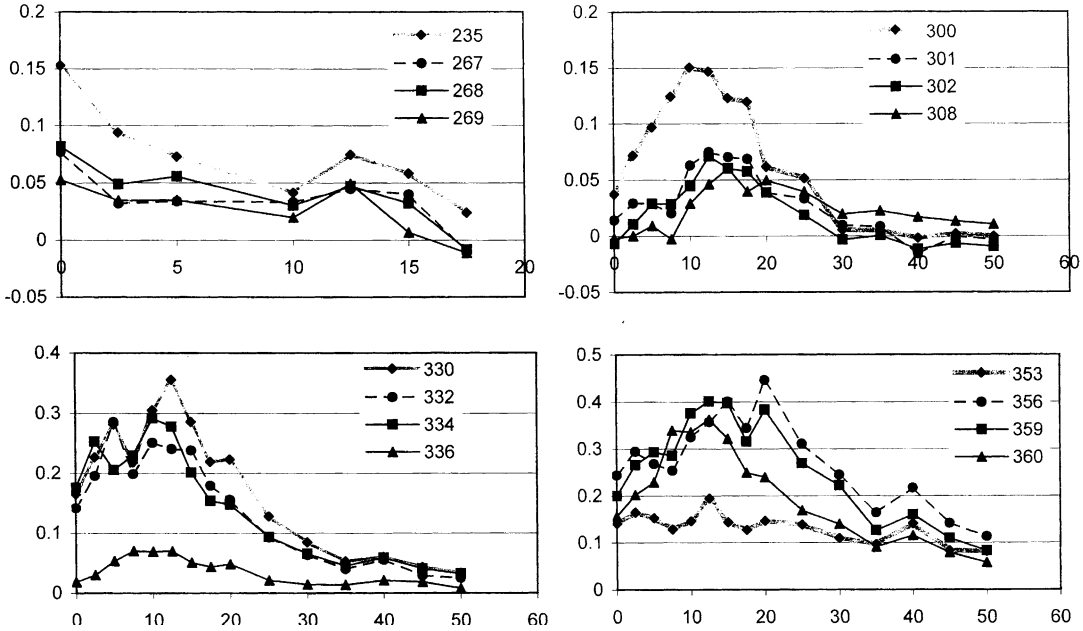

Figure 4. Footprinting plots for the RNase I experiments (optical density versus total drug concentration in μM) obtained for paromomycin interacting with Ψ-RNA. The site numbers correspond to the sequence shown in Fig. 1. Small negative intensities arise from subtraction of control-lane intensities.

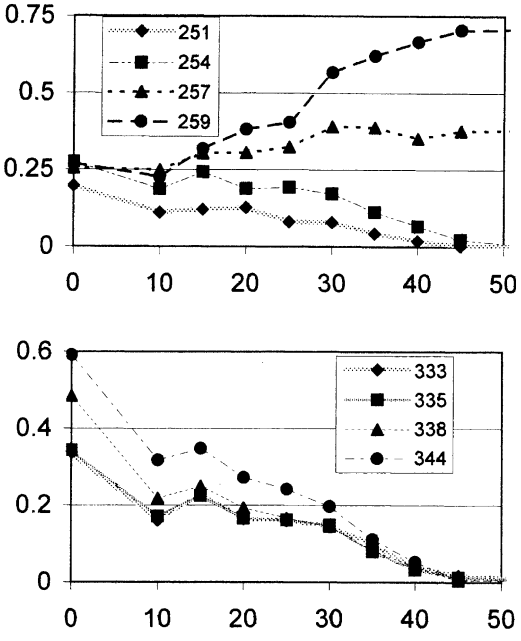

Figure 5. Footprinting plots for the RNase T1 experiments (optical density versus total drug concentration) obtained for paromomycin interacting with Ψ-RNA. The site numbers correspond to the sequence shown in Figure 1.

Table 1. Binding constants (in M^{-1}) determined from footprinting studies

Site numbers	With RNase I	With RNase T1
A235, A239	1.6×10^5	
G251, G254		2.8×10^4
C267, A268, A269	8×10^4	
A276	1.7×10^5	
A286	6×10^4	
G328, G329, G331, G333, G335		7×10^4
G338, G339, G340		6×10^4
G344, G346		7×10^4
G363, G364, G365, G366, G367		2.6×10^4

and at the pentameric guanine sequence, G363, G364, G365, G366, and G367, in the main stem. These paromomycin binding sites have values of K_i of 2.8×10^4 M^{-1} and 2.6×10^4 M^{-1}, respectively. Type 3 plots for this enzyme, for which cutting increases with drug concentration, are in the loop regions of SL-1 (G257, G259, and G261), and SL-3 (G317, G318 and G320). Spot intensities for the loop region of SL-2 were obscured by a strong band in the control lane of the gel.

Circular dichroism titrations

Circular dichroism (CD) spectra in the presence and absence of paromomycin are shown in Figure 6. The CD spectrum of the ψ-RNA resembles that for A-form RNA,[32,33] with bands at 266, 240 and 208 nm, having values of Δε (M^{-1} cm^{-1} $nucleotide^{-1}$) of +3.9, −0.8, and −3.6, respectively. With added drug, the negative band at 208 nm intensifies, and its position shifts slightly, to 210 nm. The maximum intensity is attained at ∼20 μM paromomycin concentration. At higher paromomycin concentrations, this band decreases in intensity (becomes less negative) with added drug (Fig. 6, inset). A similar but weaker effect occurs for the low-intensity band at 240 nm.

In order to determine the number of binding events and to calculate apparent paromomycin binding constants for the events, we smoothed the data of Figure 6 (using a Fourier transform algorithm) and obtained, from the smoothed data, Δε as a function of drug concentration. Some these results are plotted in Figure 7. At the higher wave lengths, the changes in Δε were too small to be analyzed. In each plot of Figure 7, Δε decreases with total drug concentration, D_T, for small D_T and levels off as D_T approaches 20 μM For D_T exceeding 20 μM, Δε increases again, possibly because another binding event that alters RNA structure occurs, or possibly because

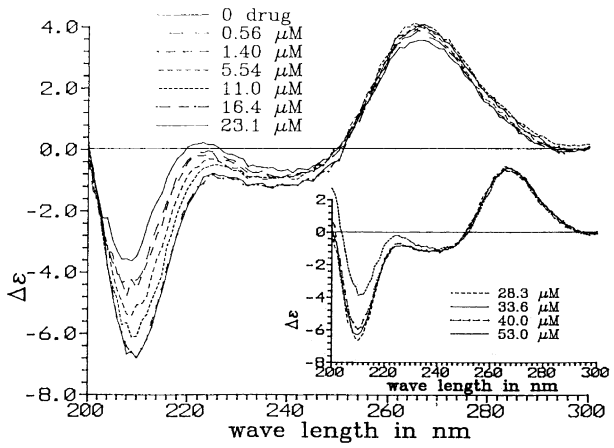

Figure 6. CD spectra of Ψ-RNA, Δε (M^{-1} cm^{-1} nucleotide^{-1}), for various paromomycin concentrations. For clarity, spectra for [paromomycin] <20 μM are shown in the main plot, and spectra for [paromomycin] >20 μM are shown in the inset. As seen in Figure 7, the intensity of the negative band increases with drug concentration below 20 μM and decreases with drug concentration above 20 μM.

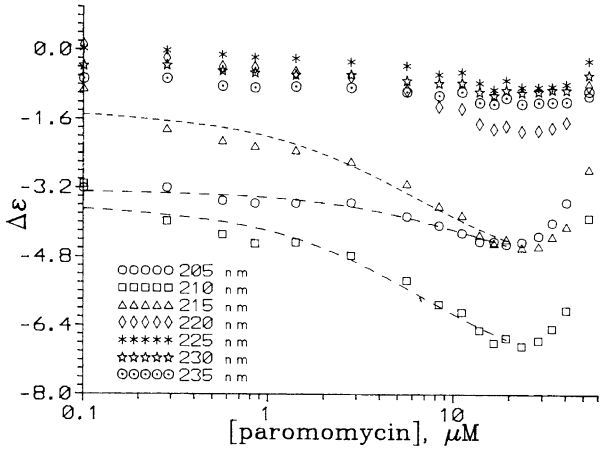

Figure 7. Observed circular dichroism, Δε (M^{-1} cm^{-1} nucleotide^{-1}), at various wave lengths, as a function of the logarithm of the drug concentration. For wave lengths above 240 nm, the changes in Δε with [paromomycin] are too small to be analyzed. For the first three wave lengths, the results of fitting to a two-state model (single drug binding event) are shown as dashed lines.

RNA is precipitating from solution. Because of the latter possibility, the event responsible for Δε becoming less negative at $D_T > 20$ μM was not analyzed. It can also be seen that Δε decreases much faster with D_T for concentrations below about 1.5 μM than it does for higher concentrations. This suggests that there are at least two CD-detectable binding events for $D_T < 20$ μM. Therefore, we analyzed the data two ways. We first obtained a single binding constant from all the data for $D_T < 20$ μM, with the understanding that this is an average or effective binding constant for the binding events taking place in this concentration range. Second, we analyzed the five data points for $D_T < 1.5$ μM to obtain an approximate value for the binding constant associated with the strongest binding event.

In each analysis, it was assumed that the change in Δε with drug concentration was due to a single drug-binding event. Suppose that, at some wave length, $\Delta\varepsilon = \alpha$ for the RNA with no drug bound and $\Delta\varepsilon = \beta$ for the RNA with drug bound. If the binding constant is K,

$$\Delta\varepsilon = \alpha\frac{1}{1+K[\mathrm{D}]} + \beta\frac{K[\mathrm{D}]}{1+K[\mathrm{D}]} = \alpha\frac{[\mathrm{R}]}{R_T} + \beta\frac{[\mathrm{RD}]}{R_T} \qquad (6)$$

Let R_T be the total RNA concentration, [R] the concentration of RNA with no drug bound, and [RD] the concentration of RNA with drug bound ($R_T = [\mathrm{R}] + [\mathrm{RD}]$). Since

$$\frac{[\mathrm{RD}]}{[\mathrm{R}][\mathrm{D}]} = \frac{[\mathrm{RD}]}{(R_T - [\mathrm{RD}])(D_T - [\mathrm{RD}])} \qquad (7)$$

is a quadratic equation for [RD], one can obtain, using eqs 6 and 7, Δε as a function of R_T, D_T, α,β and K. R_T is known, and one seeks the values of α,β and K which give the best fit to the experimental values of Δε as a function of D_T.

Some of the fits to the data for $D_T < 20$ μM are shown in Figure 7. They correspond to $K = 7.3\times10^4$ M^{-1}, 1.6×10^5 M^{-1}, and 1.7×10^5 M^{-1}. When all seven data sets shown in Figure 6 are analyzed in this way, we obtain $K = (1.28\pm0.35)\times10^5$ M^{-1} (average ± average deviation from the mean). Performing the same analysis on the data for $D_T < 1.5$ μM (Δε for the five lowest values of D_T) for wave lengths of 205, 210, 215, 220, and 225 nm (see Fig. 7), we obtain $K = (6\pm6)\times10^6$ M^{-1} (average ± average deviation from the mean). The large scatter is because most of the change in Δε takes place in the first three points. We may also estimate the uncertainty in K by seeking the changes in K (the same K is used for all wave lengths) which increase the sum of the squared deviations of calculated from experimental Δε over the best fit by 10%. This gives ln $K = 15.6\pm1.1$ (K is uncertain within a factor of 3).

Discussion

Structure of the ψ-RNA

The secondary structure of the packaging region of HIV was previously characterized using chemical and enzymatic cleavage agents and by application of folding programs to the sequence.[34–37] While there is some disagreement about the structure of certain duplex regions, most of the evidence suggests that the packaging region has four stem-loops, designated SL-1 through SL-4 in Figure 1. Clever et al.[35] showed that while SL-4 seems to be present, it contains a relatively high percentage of G·U base pairs and it is probably less stable than the other stem-loops. In addition to these four stem-loops, the ψ-RNA used in this study has sequences that extend in the 3′ direction into the coding region of the *gag* gene, and, in the 5′ direction, into the primer binding site. A

check of the sequence of the ψ-RNA used in this study revealed that T7 RNA polymerase failed to transcribe a pentameric sequence in the template DNA and that the pentamer, AAAUU (303–307), in the coding sequence of HIV-1 (LAI), is missing in the ψ-RNA. The prediction of the secondary structure of the ψ-RNA was done using Mfold v3.0 algorithm[38] offered at http://bioinfo.math.rpi.edu. Folding was done at 37 °C with 1 M NaCl, specifying that the regions between SL 3 and SL 4, and SL 4 and the main stem, be single stranded (the enzymes cut in these regions). Since the footprinting and CD studies were done at a different temperature (21 °C) and ionic strength (10 mM Tris–HCl, pH 7) from those employed in the folding program Mfold, the complete secondary structure is unknown. Thus, pending the outcome of additional work, the structure shown in Figure 1 is only intended to be a 'working model' on which the binding/structural information for paromomycin is summarized. Mfold[38] predicts that the sequences on the 3′ side of SL-4 and to the 5′ side of SL-1 form a 'main stem' (Fig. 1). This is consistent with the general lack of RNase I cutting in this region (RNase I prefers to cut at single stranded regions) (Fig. 2). Neither enzyme cuts in the linker region between SL-1 and SL-2. While this region was cut in an RNA fragment studied by Clever et al.,[35] it appears to be hidden in the ψ-RNA used in this study. The loop regions of SL-1, SL-3, and SL-4 are cut only by RNase T1. When drug binds, sites in the loops of SL-1 and SL-3 exhibit cleavage enhancements, while those in SL-4 give type 1 footprinting plots, indicative of drug binding (Figs 2–5).

The packaging region of HIV can form a dimer in which the palindromic sequence present in the loop of SL-1 is base-paired with its complement in a second copy of the genomic-length RNA. However, the fact that sites in this loop are cleaved strongly by RNase T1 indicates that the ψ-RNA is monomeric.

Binding of paromomycin to the ψ-RNA

In this work we study the binding of the aminoglycoside drug, paromomycin, to a complex RNA molecule from the packaging region of HIV-1 (LAI). The footprinting data show how drug affects cutting at specific sites. For some sites, the footprinting plots (spot intensity versus drug concentration) were type 1, indicating drug binding which blocks cleavage. Apparent binding constants for paromomycin at such sites were obtained by analyzing the footprinting plots using a two-state model, which assumes that one drug molecule was involved in the binding event. This gave values of K that were significantly lower than those obtained by analyzing, using the two-state model, the changes in the absorption[26] and CD spectra at low drug concentrations.

There could be several reasons why the values of K from footprinting disagree with those from the optical methods. (1) It is possible that the highest affinity binding site (detected by optical methods) was not measured in the footprinting experiments. Using RNase I and RNase T1, cleavage intensities high enough to be measured were obtained for ~40% of the nucleotides of the RNA. The drug would be expected to cover ~5 nucleotides and, although the enzyme may be blocked from cutting at several more sites, there are regions for which no cutting information is available. (2) The discrepancies between the values of K_i between footprinting and optical methods may be due to competitive binding in the former. If the enzyme competes with the drug for binding at the same sites, the values of K_i derived from footprinting intensities will be too low, although relative values will be correct. Since effective competition between drug and cutting agent depends in a complex way on the concentrations of drug and cutting agent, their binding constants for RNA, and the number of cutting sites for which both compete, it is difficult to prove that competition is the reason for the discrepancy. Competitive binding does explain the apparent difference between K_i values obtained from the two enzymes. However, the binding constant of one of the enzymes, RNase T1,[39] is $\sim 3\times 10^4$ M^{-1}, which is about 2 orders of magnitude lower than K_i for paromomycin binding to the ψ-RNA determined by optical methods[26] and paromomycin binding to other RNAs.[14] (3) It may be that the binding which produces the distortions detected by optical methods is different in character from that detected by footprinting. For the former, it may suffice for drug to be near the RNA molecule, without actually binding to a specific site. For bound drug to induce a decrease in cleavage, the drug molecule must block the approach of the cleavage agent to the site.

Inspection of the autoradiograms and the footprinting plots, Figures 2–5 and Table 1, reveals that the highest-affinity sites detected by RNase I are at positions A235, A239, and A276 while the sites of highest affinity detected by RNase T1 are guanine residues in SL-4: 338–340, 344, and 346. Whether the various nucleotides come together to form a binding pocket (s) for paromomycin is difficult to determine. However, it is conceivable that they could all be responding to the same drug binding event, in which case the values of K_i reported for the site are not the same for the two enzymes (Table 1). This would imply that drug and enzyme are in competition for the site, and that RNase T1 binds more strongly to RNA than RNase I (Table 1). The more likely possibility is that these nucleotides form more than one binding pocket for paromomycin and that the nucleotides affected are part of more than one site.

The RNase I footprinting results clearly show that, when A235, A239, and A276 bind drug, many enhancements (type 2 plots) occur in the single-stranded linker regions of the ψ-RNA. Although RNase T1 shows fewer enhancements (type 3 plots), they occur in the loop regions of SL-1 and SL-3, where RNase I does not cut. Since CD and absorption show that the orientation of bases and the amount of base stacking change when drug binds to the highest affinity sites, the enhancements observed in the footprinting experiments appear to be structural in origin.

Interestingly, RNase I and RNase T1 report different behavior for the homopurine tract at positions 325–336 (sequence GAAGGAGAGAGA). RNase I, which cuts

at the adenines, gives type 2 footprinting plots, but RNase T1 gives type 1 plots at the guanines in the sequence. A possible explanation for this is that the region is undergoing a structural change induced by drug binding at another site j, and that this change facilitates cutting by RNase I but hinders cutting by RNase T1. Although RNase I does not cut in SL-4 itself, the footprinting plots look more like binding plots as one approaches the 3′ end of the homopurine tract, that is, the intensity of the maximum is decreased and it is shifted to lower drug concentration. This is consistent with the nucleotides in SL-4 being part of a binding pocket for the drug.

As shown in Table 1, both enzymes report secondary binding sites on the ψ-RNA. The drug site seen by RNase T1 at G251 and G254 in the stem of SL-1 may be the same drug site seen by RNase I at C267, A268, and A269 on the other 'side' of the stem. If this is the case, the difference in the values of K_i by about a factor of three may be explained by competitive binding, with RNase T1 being the stronger binder of the two enzymes toward RNA. RNase T1 also detects a weaker site at an oligo-guanine tract at 363–367, which is located in the main stem of the RNA.

From the CD changes (Figs 6 and 7) it is clear that paromomycin changes the structure of the ψ-RNA. Since the intensity of the 208 nm band is affected when the linker regions and the loops of SL-1 and SL-3 are undergoing footprinting enhancements, this CD band is affected by the relative orientation of bases in these regions. As discussed above, CD detects a high-affinity binding event for drug concentrations below 1.5 μM, and another event in the range 1.5–20 μM. Interpreting both according to the two-state model, we find effective binding constants of $(6\pm6)\times10^6$ M^{-1} and $(1.28\pm0.35)\times10^5$ M^{-1}. It is possible that the change in CD in each range is due to multiple drug-binding events with the same or similar values of K_i, each of which changes the CD in a different way. This situation would require analysis using a model more complex than that employed. Many of the footprinting plots obtained with RNase I show enhancement followed by a decrease in cutting at higher drug concentration, type 2 plots. The increasing parts of these plots can be explained by a drug-induced structural change while the decreasing portions could be due to either a second structural change or drug binding. Determining which situation is the case will require additional study.

The packaging region of HIV initiates the formation of the genomic length RNA dimer found in the virion of the virus.[25] This is believed to occur through SL-1, which contains the palindromic sequence GCGCGC at positions 257–262 of the ψ-RNA.[40,41] The RNase T1 footprinting studies clearly show that this loop is free, and that it undergoes a structural change in the presence of paromomycin. To what extent paromomycin may affect the monomer-dimer equilibrium involving the ψ-RNA remains to be seen. However, the bulge region of SL-1, adjacent to A276 and part of a binding pocket for the drug, is thought to regulate the inter-conversion between the 'kissing' and extended duplex forms of the HIV RNA dimer.[42] The packaging region is also bound by the nucleocapsid protein NCp7 in the virion of HIV.[25] Since drug binding could displace protein from RNA, paromomycin and its analogues could affect the organization and compaction of RNA in the virus particle which could in turn influence viral infectivity.

Conclusions

In this work, we use quantitative footprinting methods to identify binding sites for the aminoglycoside drug paromomycin on a segment of RNA from the packaging region of HIV-1 (LAI). The footprinting autoradiographic data were used to construct plots (called footprinting plots) showing how cutting at various sites changes with drug concentration. Plots showing decreases in spot intensity with drug concentration are referred to as type 1, or binding plots. These were analyzed using a two-state binding model, assuming a binding stoichiometry of 1:1, to give apparent binding constants for the drug. The values of K derived using this approach were lower than the values of K obtained from changes in the CD spectrum of the RNA at low drug concentration. The highest affinity paromomycin site detected by RNase I involves nucleotides in SL-1 and the main stem and it may be part of the same site detected by RNase T1 in SL-4. Footprinting and CD show that drug binding to this site causes the linker regions and loops of SL-1 and SL-3 to change conformation. These studies show that quantitative footprinting in conjunction with circular dichroism can provide information concerning drug binding sites and regions of structural change on a segment of RNA from the packaging region of HIV.

Experimental

Footprinting

The synthesis, purification and 5′-end labeling (^{32}P) of the 171-mer ψ-RNA were given earlier.[25] Preliminary experiments were necessary to determine the amounts of RNase I and RNase T1 and the cutting time required for 'single hit' conditions, such that less than ~20% of the full length RNA is cut during the reaction. Enzymes were chosen over a number of possible chemical cleavage agents because the extent of digest with an enzyme is easily controlled (required for quantitative studies) and the end-chemistry at the site of cleavage is homogeneous (i.e., one oligomer produced by cutting corresponds to one nucleotide position on the polymer).

Footprinting experiments with paromomycin were carried out in a total volume of 10 μL in the buffer, 10 mM Tris HCl (pH 7), at room temperature (~20 °C). Under these conditions, the RNA is in the monomeric form. The protocol was to mix radiolabeled RNA (100,000 cpm) with unlabeled RNA, and to add increasing amounts of paromomycin to successive samples in a total volume of 9 μL. After mixing with a pipette, each

solution was allowed to stand for 30 min. Then RNase I or RNase T1 in a volume of 1μL was added and the sample briefly mixed. The final concentrations of labeled RNA and unlabeled RNA in the final reaction volume of 10 μL were ~0.05 and 1.1 μM, respectively. For the RNase I experiments, the final drug concentrations were: 0, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20, 25, 30, 35, 40, 45, and 50 μM. For the RNase T1 experiments, they were: 0, 10, 15, 20, 25, 30, 35, 40, 45, 55, and 60 μM.

Cutting was allowed to proceed for 1.5 min (RNase I) or 10 min (RNase T1) and the cleavage reactions were stopped by addition of 5 μL formamide loading buffer. Before loading in the gel, samples were heated at 95 °C for 5 min and then immediately placed on ice. Eleven μL of each sample was loaded into an 9% denaturing PAGE gel, which ran at 800 V for 8 h at room temperature. The cutting pattern was visualized with autoradiography (Figs 2 and 3) and the data were scanned using the earlier described procedures.[24] The spot intensities in the control lanes show that a small amount of the end-labeled ψ-RNA was fragmented prior to addition of the cleavage agent; sites having a significant amount of cutting in the control were not used in the quantitative analysis. For sites showing binding, band intensity for the fully occupied site was taken as the average of the intensity at high drug concentration. The corrected spot intensities are used to construct footprinting plots, some of which are shown in Figs 4 and 5.

Circular dichroism

Circular dichroism titrations were carried out at 21 °C using an Aviv model 202 CD instrument. Scans of wavelength were from 300 to 200 nm with a resolution of 1 nm and data sampling every 2 nm. The starting volume in the 1 cm path length cell was 350 μL with the concentration of RNA being 0.5 μM in the buffer, 10 mM Tris–HCl (pH 7.0). Control experiments, in which spectra were collected at various times after addition of paromomycin to RNA and mixing, showed that binding equilibrium was reached within the time of mixing and collecting the data.

Spectra were collected immediately after addition of a small volume of a solution of paromomycin (in the above buffer) to a buffered solution of the RNA. After subtracting the spectrum of the buffer blank, the spectra were corrected for dilution (correction less than 10% at the end of the titration) and $\Delta\varepsilon$ per nucleotide was calculated. The formula used was $\Delta\varepsilon = \psi/[(32.98)(171)\ell c]$, where $\Delta\varepsilon$ is in $M^{-1}\,cm^{-1}$, ψ is the measured ellipticity in degrees, l is the path length in cm, and c is the concentration of RNA, 0.5 μM. The paromomycin concentrations used in the titration were: 0, 0.28, 0.56, 0.76, 1.03, 1.40, 2.78, 5.54, 8.28, 11.0, 13.7, 16.4, 19.1, 23.1, 28.6, 33.6, 40.1, and 53.0 μ. Results for the first six concentrations were used to determine the large binding constant. Paromomycin alone does not exhibit a CD spectrum in the wavelength range that was studied.

Acknowledgements

We wish to thank Professor Y. Tor (U. of California at San Diego) and P. N. Borer (Syracuse University) for helpful discussions. We also wish to thank Professor S. Loh (SUNY Upstate Medical University) for the use of the CD instrument. The research was supported in part by the NIH (GM32691 to P. N. Borer) and the Chemistry Department.

References and Notes

1. Moazad, D.; Noller, H. F. *Nature* **1987**, *327*, 389.
2. Schroeder, R.; Waldsich, C.; Wank, H. *EMBO J*. **2000**, *19*, 1.
3. Recht, M. I.; Fourmy, D.; Blanchard, S. C.; Dahlquist, K. D.; Puglisi, J. D. *J. Mol. Biol*. **1996**, *262*, 421.
4. Recht, M. I.; Douthwaite, S.; Dahlquist, K. D.; Puglisi, J. D. *J. Mol. Biol*. **1999**, *286*, 33.
5. Fourmy, D.; Yoshizawa, S.; Puglisi, J. *D. J. Mol. Biol*. **1998**, *27*, 333.
6. Vicens, Q.; Westhof, E. *Structure* **2001**, *9*, 647.
7. Jiang, L.; Majumdar, A.; Hu, W.; Jaishree, T. J.; Xu, W.; Patel, D. *J. Struct. Fold. Des*. **1999**, *7*, 817.
8. Varani, L.; Spillantini, M. G.; Goedert, M.; Varani, G. *Nucleic Acids Res*. **2000**, *28*, 710.
9. Faber, C.; Sticht, H.; Schweimer, K.; Rösch, P. *J. Biol. Chem*. **2000**, *275*, 20660.
10. Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630.
11. Tok, J. B.-H.; Huffman, G. R. *Bioorg. Med. Chem. Lett*. **2000**, *10*, 1593.
12. Tok, J. B.-H.; Dunn, L. J.; Des Jean, R. C. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 1127.
13. Cho, J.; Rando, R. R. *Biochemistry* **1999**, *38*, 8548.
14. Ryu, D. H.; Rando, R. R. *Bioorg. Med. Chem*. **2001**, *9*, 2601.
15. Llano-Sotelo, B.; Chow, C. S. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 213.
16. Wang, S.; Huber, P. W.; Cui, M.; Czarnik, A. W.; Mei, H.-Y. *Biochemistry* **1998**, *37*, 5549.
17. Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. *Proc. Natl. Acad. Sci. U.S.A*. **1999**, *90*, 10129.
18. Litovchick, A.; Evdokimov, A. G.; Lapidot, A. *Biochemistry* **2000**, *39*, 2838.
19. Kirk, S. R.; Luedtke, N. W.; Tor, Y. *J. Am. Chem. Soc*. **2000**, *122*, 980.
20. Wang, H.; Tor, Y. *Bioorg. Med. Chem. Lett*. **1998**, *8*, 3665.
21. Michael, K.; Wang, H.; Tor, Y. *Bioorg. Med. Chem. Lett*. **1999**, *7*, 1361.
22. Baker, T. J.; Luedtke, N. W.; Tor, Y.; Goodman, M. *J. Org. Chem*. **2000**, *65*, 9054.
23. Hermann, T. *Angew. Chem., Int. Ed*. **2000**, *39*, 1890.
24. McPike, M. P.; Goodisman, J.; Dabrowiak, J. C. *Methods Enzymol*. **2001**, *340*, 431.
25. Coffin, J. In *RNA Tumor Viruses*; Weiss, R., Teich, N., Varmus, H., Coffin, J., Eds.; Cold Spring Harbor Laboratory: Cold Spring Harbor, 1985; Vol 2, pp 17–74.
26. Sullivan, J. M.; Goodisman, J.; Dabrowiak, J. *C. Bioorg. Med. Chem. Lett*. **2002**, *12*, 615.
27. Dabrowiak, J. C.; Goodisman, J.; Ward, B. In *Drug-DNA Interaction Protocols*; Fox, K. R., Ed.; Humana: Totowa, 1997; p 23.
28. Dabrowiak, J. C.; Stankus, A. A.; Goodisman, J. In *Nucleic Acid Targeted Drug Design*; Probst, C. L., Perun, T. J., Eds.; Marcel Dekker: New York, 1992; p 93.

29. D'Alessio, D.; Riordan, J. F. *Ribonucleases. Structures and Functions*; Academic: San Diego, 1997.
30. Trauger, J. W.; Dervan, P. B. *Methods Enzymol*. **2001**, *340*, 450.
31. Stankus, A.; Goodisman, J.; Dabrowiak, J. C. *Biochemistry* **1992**, *31*, 9310.
32. Johnson, C. W. *Methods Biochem Anal*. **1985**, *31*, 61.
33. Rigl, C. T.; Lloyd, D. H.; Tsou, D. S.; Gryaznov, S. M.; Wilson, D. W. *Biochemistry* **1997**, *36*, 650.
34. Berkhout, B. *Prog. Nucl. Acids Res*. **1996**, *54*, 1.
35. Clever, J.; Sassetti, C.; Parslow, T. G. *J. Virol*. **1995**, *69*, 2101.
36. McBride, M. S.; Panganiban, A. T. *J. Virol*. **1996**, *70*, 2963.
37. Zeffman, A.; Hassard, S.; Varani, G.; Lever, A. *J. Mol. Biol*. **2000**, *297*, 877.
38. Mathews, D. H.; Sabrina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol*. **1991**, *288*, 911.
39. Steyaert, J. *Eur. J. Biochem*. **1997**, *247*, 1.
40. Langhrea, M.; Jetté, L. *Biochemistry* **1997**, *36*, 9501.
41. Mariaux, D.; De Rocquigny, H.; Roques, B. P.; Paoletti, J. *J. Biol. Chem*. **1996**, *270*, 33686.
42. Takahashi, K.; Baba, S.; Hayashi, Y.; Koyanagi, Y.; Yamamoto, N.; Takaku, H.; Kawai, G. *RNA* **2000**, *6*, 96.