Oxidative base damage in RNA detected by reverse transcriptase

Youngsook Rhee, Michael R. Valentine and John Ternini*

Department of Molecular Biochemistry, Beckman Research Institute of the City of Hope, 1450 E. Duarte Road, Duarte, CA 91010, USA

Received April 17, 1995; Revised and Accepted July 12, 1995

ABSTRACT

Oxidative base damage in DNA and metabolic defects in the recognition and removal of such damage play important roles in mutagenesis and human disease. The extent to which cellular RNA is a substrate for oxidative damage and the possible biological consequences of RNA base oxidation, however, remain largely unexplored. Since oxidatively modified RNA may contribute to the high mutability of retroviral genomic DNA, we have been interested in developing methods for the sequence specific detection of such damage. We show here that a primer extension assay using AMV reverse transcriptase (RT) can be used to reveal oxidatively damaged sites in RNA. This finding extends the currently known range of RNA modifications detectable with AMV reverse transcriptase. Analogous assays using DNA polymerases to detect base damage in DNA substrates appear to be restricted to lesions at thymine. Oxidative base damage in the absence of any detectable chain breaks was produced by dye photosensitization of RNA. Six out of 20 dyes examined were capable of producing RT detectable lesions. RT stops were seen predominantly at purines, although many pyrimidine sites were also detected. Dye specific photofootprints revealed by RT analysis suggest differential dye binding to the RNA substrate. Some of the photoreactive dyes described here may have potential utility in RNA structural analysis, particularly in the identification of stem-loop regions in complex RNAs.

INTRODUCTION

Oxidative base damage of DNA by endogenous and environmental sources plays an important role in mutagenesis. To date, more than 30 modified bases produced by radical mediated oxidative damage to DNA have been described (1), but the promutagenic consequences of only a small subset have been explored. Although one might imagine that the nuclear chromatin structure offers some degree of protection against oxidative DNA damage, it has been estimated that the genomic DNA of each cell receives ~10 000 oxidative hits per day (2). Failure to remove these lesions from DNA can result in the introduction of hereditable mutations via formation of stable mispairs during translesional synthesis.

In contrast, little is known about the extent and distribution of oxidative base damage in RNA. RNA may be considered to have enhanced potential for oxidative attack due to its widespread cytosolic distribution within various organelles. Oxidatively damaged RNA can interfere with correct base pairing, which would compromise the accuracy of cellular processes such as transcription and translation. Where RNA serves as the genomic material, the potential for mutagenesis exists as a result of base misincorporation opposite oxidatively damaged templates. In the case of retroviral replication, base damage to RNA can contribute to an elevated mutation rate in DNA.

One of the promutagenic mispairs in DNA resulting from guanine oxidation followed by translesional synthesis, 8-oxAdG: dA, has been well characterized structurally (3,4). The mispair gives rise to G → T transversions in both procaryotes (5–7) and eucaryotes (8,9). 8-oxoguanosine has recently been identified in RNA (10) as have other oxidized bases previously identified in DNA such as 8-hydroxyadenine, 5-hydroxycytosine and 5-hydroxymethyluridine (11). The presence of these oxidized bases, coupled with a lack of known repair processes for RNA and the absence of proofreading activities for the various enzymes which copy RNA templates, indicate the potential for mutagenesis by oxidatively damaged RNA. In order to examine the scope of oxidative base damage in RNA and its possible biological consequences, sequence specific detection methods are required.

In the case of DNA, both enzymatic and chemical approaches have been used for the sequence specific detection of base damaged sites. Repair glycosylases sensitive to various types of oxidative DNA lesions have been used in conjunction with DNA sequencing methods to locate base damage (12,13). Some, like endonuclease III (nth protein), recognize a class of lesion (dihydropyrimidine) rather than a specific chemical modification. Only a limited number of repair glycosylases have been characterized to date and fewer are readily available in the quantities required for use as sequencing reagents. It also appears unlikely that an arsenal of glycosylases will be available for every oxidative lesion, thus the universal applicability of this approach is unclear.

The presence of certain types of base lesions in template DNA can be detected by assays utilizing primer extension with DNA polymerase. Encountering these lesions during DNA synthesis, the polymerase is unable to ‘bypass’ the site to any appreciable

* To whom correspondence should be addressed
extent and the pausing at such sites gives rise to aborted products whose size corresponds to termination either at the damaged site or the nucleotide immediately preceeding it. In this way, pyrimidine dimers and bulky base adducts resulting from reaction of DNA with aminofluorene, psoralen (14), cis-platinum (15) and benzo[a]pyrene diol epoxide (16) can be detected. For the detection of oxidative base damage however, the generality of this approach is unclear. The thymine glycol oxidative lesion has been thoroughly examined with respect to its ability to inhibit the processivity of various DNA polymerases. T4 DNA polymerase is quantitatively arrested opposite this lesion (17), whereas the ratio of pausing:bypass for other polymerases appears to be a complex function of sequence, template length and 3', 5' exonuclease activity (18-20). The presence of 8-oxodG in the template strand does not appear to impede chain extension by DNA polymerases (6) therefore, it appears unlikely that primer extension methods can be used to detect this lesion.

Amine nucleophiles which catalyze the elimination of the oxidized base and promote subsequent chain cleavage have also been used in cases where the oxidative base modification results in a weakening of the C1'-N glycosidic bond (21-23). Some oxidative base modifications do not appreciably weaken the C1'-N bond strength relative to unmodified bases; thus, they would not be labilized by amine nucleophiles such as piperidine. The oxidative base lesions 5-(hydroxymethyl) deoxyuridine(5-hMdU) and 4,8 dihydro-4-hydroxy-8-oxo-deoxyguanosine, are in this category (24).

Since there are no existing methods for the sequence specific detection of oxidative base damage in RNA, several approaches analogous to those described for DNA damage were explored. No glycosylase activities have been identified which remove oxidatively damaged bases in RNA and our attempts to use the 8-oxodG:dC specific glycosylase (fpg protein) to induce chain breaks at the analogous RNA lesion sites failed (Rhee and Termini, unpublished). Aniline has been used in place of piperidine for the chemical labilization of modified bases in RNA, owing to its decreased basicity relative to piperidine (25). In addition to the limitations described above for non-labile lesions, backbone cleavage due to 2'OH mediated phosphodiester hydrolysis is often problematic in RNA, obscuring cleavages arising from oxidative base loss.

We report here a novel method using an AMV reverse transcriptase (RT) primer extension assay for the sequence specific mapping of oxidative base damage in RNA. Pausing produced by oxidative damage at all bases in RNA can be detected simultaneously in a single primer extension run. This finding extends the range of RNA substrate modifications which can impede DNA primer extension catalyzed by RT. Reverse transcriptase has previously been shown to stop at DMS and kethoxal modified bases (26), crosslinks (27), psoralen adducts (28); as well as 'bulky' hypermodified bases (29). Reverse transcriptase has also been used to detect hypermodified bases in rRNA (30), and is an indispensable reagent for the structural analysis of complex RNA molecules based on the interruption of cDNA synthesis at chemically modified bases or unusually stable secondary structures (31).

In order to provide reference oxidative base damage, dye photosensitization methods were used. Photooxidation of nucleic acid substrates using visible light absorbing dyes such as methylene blue is a convenient method for the production of base damage in the absence of chain breaks (32). It has been shown that base modifications at guanosine in DNA resulting from dye photosensitization methods are identical to those produced by endogenous processes in vivo (33). The recent report of 8-oxoguanosine formation in RNA via methylene blue and rose bengal photosensitization (10) provided additional impetus for the use of this methodology to create oxidative base damage in RNA.

Primer extension analyses with AMV-RT following dye photosensitization reveal dye specific photofootprints. Thus the RT assay has allowed us to differentiate patterns of dye binding based on the oxidative footprints obtained subsequent to equilibrium binding and photolysis. Single stranded loops of stem--loops are not substrates for photooxidation, while double helical regions appear to be preferred oxidation sites. Some if not all of the variability observed in the photofootprints arise from dye-RNA binding interactions. Nile blue and riboflavin display marked GC specificity, as determined by RT photofootprinting and homopolymer binding studies. Scatchard analysis was used to confirm a relatively strong affinity of thionin (Ka = 6x10^6 M^-1) to a model RNA substrate. Thus the RT assay has allowed us to discriminate between different patterns of dye binding based on the oxidative footprints obtained subsequent to equilibrium binding and photolysis. More importantly, primer extension studies have shown that AMV reverse transcriptase stops at a variety of oxidized purine and pyrimidine bases, indicating its use as a probe of endogenous oxidative base damage of RNA.

MATERIALS AND METHODS

General assay procedures

An 80 nucleotide RNA possessing a simple secondary structure was used as a substrate in these experiments (see below). Following photosensitization of RNA, unreacted and reduced (leuco) dyes were removed by ethanol precipitation and gel filtration. The RT reactions were carried out at elevated temperatures in order to disrupt RNA secondary structures which are not easily displaced by the advancing polymerase. The synthesis of a radiolabeled cDNA from the photomodified template RNA by RT is interrupted at specific sites as the enzyme pauses or stops at the lesions. These sites were then mapped by comparison with dideoxy sequencing lanes of unirradiated control RNA.

Dyes

All dyes tested were purchased from Aldrich Chemical Co. Dye purity was 90% or greater and was confirmed by TLC analysis. Dye stock solutions (1 mM, 10x) were prepared in 1 x reverse transcriptase buffer and stored frozen in the dark until use. Concentrations were determined by visible absorption spectroscopy using calculated molar extinction values determined in the appropriate buffer. Structures and absorbance spectra for dyes found to be RNA photoreactive in these studies are provided in Figure 1. A partial list of dyes screened for photoreactivity is given in Table 1.

Preparation of RNA substrate

Plasmid pGEM9Zf(-) (50 µg, Promega), containing both the Sp6 and T7 RNA polymerase promoters, was linearized with 100 U of NotI restriction enzyme, added in two portions at 1.5 h intervals. Digestion was complete as judged by electrophoresis on 1% agarose. An aliquot of 1 µg of linearized plasmid was required
Figure 1. Visible absorbance spectra and structures of the RNA modifying dyes identified in this study. (A) Nile blue chloride, 9.5 μM; (B) napthol blue black, 12.3 μM; (C) thionin, 23.6 μM; (D) eosin Y, 9.2 μM; (E) safranin O, 10.2 μM; (F) riboflavin, 29.8 μM.

for the synthesis of 2 nmol of RNA using the T7 Ambion Megascript® kit. RNA was purified by 15% PAGE and the 80 nucleotide substrate RNA was visualized by UV shadowing and excised from the gel. RNA was recovered from the polyacrylamide by electroelution in TBE using an Elutrap® apparatus (Schleicher & Schuell). Samples were desalted by centrifugation over 1 ml of Sephadex G-25 (125 μl/ml G-25) and precipitated by adding 1/10 vol of 2.5 M NaOAc and 3 vol of EtOH at -20°C. Molar extinction coefficients were calculated using the method of Borer (34). The sequence and secondary structure of the 80mer RNA [pGEM(l-80)] are shown in Figure 2.

Photolysis conditions

RNA substrate (0.2 μg) was heated to 80°C in 5 μl of RT buffer and allowed to cool to room temperature gradually in a heat block. Aliquots of the various dye stock solutions were added to make a final dye concentration of 0.1 mM in 50 μl. After 1 h of equilibration at room temperature, samples were irradiated with light from an Oriel 1000 W Hg (Xe) arc lamp. Light was reflected using a 90° beam turner outfitted with a dichroic (cold) mirror, which reflects light from 420-630 nm with negligible throughput loss. An additional 350 nm longpass filter was used in order to remove the small amount of UV light that was reflected by the dichroic mirror. Light output was measured using an Oriel Merlin® detection system equipped with a pyroelectric detector head mounted with a CaF2 window. Irradiations were carried out for 10 min, with an output of 1.2 J cm⁻² measured at 540 nm.

RT assay

The dideoxy sequencing and primer extension analyses were adapted from Inoue and Cech (35). A DNA primer complementary to the first 15 bases of 80mer RNA substrate (Fig. 2) was endlabeled to a specific activity of ~30 000 c.p.m./μl with T4 kinase. The volume of the photolysis reactions was brought up to 100 μl and unreacted dye was removed by spin filtration through 1 μl of Sephadex G-25 prequillibrated with H2O. Ethanol precipitation was carried out as described above and the presence of residual dye was assayed by UV. Dried samples were resuspended in 5 μl of H2O. Annealing was carried out by the addition of 1 μl of 10× RT buffer (~Mg), 2 μl of [5'-32P]cDNA primer, 2 μl of H2O and heating to 80°C for 3 min followed by...
slow cooling in a heat block to 37°C. 2 µl of 35 mM Mg(OAc)$_2$ solution was added. 10 mM stock solutions of dNTPs were prepared in ddH$_2$O buffered to pH = 7.0 with 0.25 M Tris base. A 5 × stock solution of dNTPs was prepared by adding 15 µl each of 10 mM dNTP stock to 12 µl of H$_2$O and 8 µl of 10 × RT (+Mg) buffer (10 × RT buffer: 0.5 M Tris-HCl, pH = 8.3, 0.6 M NaCl, 0.06 M Mg(OAc)$_2$, 0.1 M DT) Separate ddNTP stock solutions for sequencing were prepared by the addition of 2.5 µl each of 10 mM ddNTP stocks to 80 µl of H$_2$O and 10 µl of 10 × RT(–Mg) buffer. Reverse transcriptase mix was prepared by the addition of 2 µl of 20 U/ml AMV RT in 8 µl of 1 × RT(+Mg) buffer. For dideoxy sequencing of unreacted RNA, annealing was carried out as described above; 2 µl of 35 mM Mg(OAc)$_2$ solution were added to make a 12 µl hybridization mix. To this was added 1 µl of 5 × ddNTP solution, 1 µl of 5 × dNTP stock and 1 µl of reverse transcriptase mix. For analysis of the photooxidized RNA, 2 µl of 5 × dNTP stock was added to 12 µl of hybridization mix and 1 µl of RT mix was added to commence the reaction. The reactions were carried out at 50–55°C for 30 min and were stopped by the addition of 10 µl of formamide loading buffer containing 95% formic acid, 20 mM EDTA and 0.05% BPB. Samples were heated to 90°C for 2 min and quick chilled on ice prior to electrophoresis on a denaturing (7 M urea) 10% polyacrylamide sequencing gel. Gels were transferred to Whatman paper, dried down and subjected to autoradiography on Kodak X-Omat film.

**Equilibrium dialysis**

Dialysis experiments were conducted in a Hoffer EMD101 microdialyzer using a total volume of 200 µl/chamber. RNA concentration was fixed at 6.1 µM in 10 mM MES buffer, pH = 7.0, while initial dye concentrations varied from 5.8 to 29.5 µM. The MW cutoff of the EMD103 membranes was 12 000–14 000 D. There was little or no non-specific dye binding to membrane observed over the concentration range studied for thionin. 100 µl aliquots were removed after 5–6 h equilibration time (equilibration was judged >90% complete after 3 h) and diluted 1:1 with a solution of 0.5% AcOH in DMSO. This dissociates bound dye from the RNA permitting spectrophotometric determination of complexed ligand (36); that this resulted in complete dissociation of bound thionin could be shown by a shift in λ$_{max}$ and absorbance for a 4:1 RNA:thionin solution in MES. Bound thionin absorbs maximally at 617 nm and upon addition of DMSO/0.5% AcOH this value shifts to that of free dye, 600 nm). Extinction coefficients of dyes were determined in 1:1 MES buffer: 0.5% AcOH in DMSO; and the value for thionin was calculated as 48 251 M$^{-1}$ cm$^{-1}$. Beers law was found to hold for absorbance values <1.0, implying an absence of significant dye aggregation over this concentration range.

**Fluorescence measurements**

Spectra were recorded on a Hitachi F-2000 fluorescence spectrophotometer. Stock solutions of dye and RNA were made up in 50 mM Tris–HCl, pH = 6.8. RNA homopolymers poly(A)poly(U) and poly(G)poly(C) were purchased from Sigma. Dye stock solutions were made up immediately prior to fluorescence measurements. Homopolymer RNAs of various concentrations (in excess of dye), determined as molar base pairs, were added to dye stock solutions and allowed to equilibrate for several minutes at 37°C prior to fluorescence measurements. Spectra were recorded in a thermostatted cell held at 37°C.

**RESULTS**

**RT primer extension assay reveals a subset of RNA photooxidizing agents**

Six dyes out of 20 screened were found to be capable of photomodifying RNA upon irradiation with visible light. Gel electrophoresis on 15% denaturing (7 M urea) PAGE of 5' end-labeled pGEM92F(–) 80mer RNA following photolyisis with light of wavelength >420 nm revealed an absence of chain breaks (Fig. 3, lane 6) and some minor crosslinked products (<5%, data not shown). Visualization of the photomodified sites using aniline blue染色 reveals a subset of RNA oxidation with various dyes. Lanes 1–4, C, T, A, G dideoxy sequencing lanes respectively; Lane 5, primer extension of the substrate RNA in the absence of dye treatment and photolysis; Lane 6, RNA irradiated in the absence of dyes prior to primer extension; Lanes 7–14, 80mer substrate RNA irradiated in the presence of eosin Y (EY), mide blue chloride (NB), riboflavin (RF), thionin (Th), safranin O (SO), indigo carmine (IC), naphtol blue black (NBB) respectively, as described in Materials and Methods.

Figure 3. Primer extension analysis of pGEM(1–80) RNA following photooxidation with various dyes. Lanes 1–4, C, T, A, G dideoxy sequencing lanes respectively; Lane 5, primer extension of the substrate RNA in the absence of dye treatment and photolysis; Lane 6, RNA irradiated in the absence of dyes prior to primer extension; Lanes 7–14, 80mer substrate RNA irradiated in the presence of eosin Y (EY), mide blue chloride (NB), riboflavin (RF), thionin (Th), safranin O (SO), indigo carmine (IC), naphtol blue black (NBB) respectively, as described in Materials and Methods.
observed at cytosine than at uracil. Eosin Y, nile blue chloride and loss.

RNA substrate (G19A21C19U21), the distribution of lesions at G, all four residues occur in approximately equal amounts for this napthol blue black fail to photosensitize any cytosine sites. Since photosensitization reactions, where adenine modifications out-

An exception to this trend is observed for the napthol blue black modification, with slightly more guanine than adenine stop sites. It is clear that transcriptional pausing occurs at residues other than G, implying photomodification at all four bases. RNA base assignments in Table 1 reflect a preponderance of purine G, implying photomodification at all four bases. RNA base numbering corresponds to cDNA and RNA sequences shown in Figure 2. Sites of RNA modification are shown in parentheses. ND, not detected.

<table>
<thead>
<tr>
<th>cDNA (RNA)</th>
<th>Eosin Y</th>
<th>Nile blue chloride</th>
<th>Riboflavin</th>
<th>Thionin</th>
<th>Celestine blue</th>
<th>Safranin O</th>
<th>Indigo carmine</th>
<th>Napthol blue black</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(U)</td>
<td>20.37,57</td>
<td>18.20,37,57</td>
<td>20.37,57</td>
<td>18.20,37,40,37,57</td>
<td>ND</td>
<td>20.37,57</td>
<td>ND</td>
<td>18.20,37</td>
</tr>
<tr>
<td></td>
<td>(61,44,24)</td>
<td>(63,61,44,24)</td>
<td>(61,44,24)</td>
<td>(63,61,44,38,24)</td>
<td>ND</td>
<td>(61,44,24)</td>
<td>ND</td>
<td>(63,61,44)</td>
</tr>
<tr>
<td>G(C)</td>
<td>ND</td>
<td>ND</td>
<td>32.45,56</td>
<td>32.45,50</td>
<td>ND</td>
<td>45.50</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(49,36,25)</td>
<td>(49,36,31)</td>
<td></td>
<td>(36,31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(G)</td>
<td>19.21,36,38,58</td>
<td>19.36,38,58</td>
<td>19.21,33,36,58</td>
<td>19.21,36,58</td>
<td>ND</td>
<td>19.21,36,38,58</td>
<td>ND</td>
<td>19.36,66</td>
</tr>
<tr>
<td></td>
<td>46.51,53,58</td>
<td>(62,45,43,23)</td>
<td>38.51,53,58</td>
<td>38.46,58</td>
<td></td>
<td>46.51,58</td>
<td>(62,45,15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(62,60,45,43,23)</td>
<td>(62,60,48,45,23)</td>
<td>(62,60,48,45,23)</td>
<td>(62,60,45,43,23)</td>
<td>(62,60,45,43,23)</td>
<td>(62,60,45,43,23)</td>
<td>(62,60,45,43,23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35,30,32,23</td>
<td>(43,30,28,23)</td>
<td>43,30,28,23</td>
<td>43,35,23</td>
<td></td>
<td>35,30,23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T(A)</td>
<td>22,39,47,59,46,22</td>
<td>35,39,59</td>
<td>22,35,39,52</td>
<td>35,39,42,52</td>
<td>ND</td>
<td>22,35,39,52</td>
<td>ND</td>
<td>39,42,49,59,22</td>
</tr>
<tr>
<td></td>
<td>(59,42,34,29)</td>
<td>(46,42,29)</td>
<td>(46,42,39)</td>
<td>48,49,59</td>
<td></td>
<td>49,52</td>
<td>(38,39,33,22)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>29,27,22</td>
<td></td>
<td>32,33,22</td>
<td>32,29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Most of the dyes tested in this assay did not yield RT detectable photolesions. Triphenyl methane dyes such as malachite green and crystal violet, indigoid dyes such as indigo carmine and several cyanine dyes such as merocyanine 540 failed to provide photofootprints. Assignment of lesions within the RNA corresponding to G6-U53 were made by comparison to dideoxy-sequencing lanes of the corresponding cDNA nucleotides A18-C76 (Fig. 3, lanes 1–4). Nucleotides C67-C80 of RNA are hybridized to primer, so any lesions within this 3' segment remain undetected. Three bands near the top of the autoradiograph are observed in unirradiated controls as well as in reactions with unreactive dyes (lanes 5, 11 and 13). These represent pauses in the cDNA synthesis resulting from structural blockades in the RNA. Stops can be assigned at A22, G15 and G43-C49. The C49-G48 dinucleotide immediately to the 5' side of this loop is only modified by thionin and riboflavin. The 5' loop corresponding to nucleotides G15–A22 likewise appears unreactive towards photomodification. Similarly, we find oxidation occurring in the stem flanking the 3' side of this loop, corresponding to bases G33–C25 of the three base pair stem. All photoreactive dyes appear to modify this stem to some extent. The riboflavin cleavage pattern in this region is displaced slightly to the 3' side relative to the corresponding bands observed with other dyes. Napthol blue black only appears to modify G23 of this stem (Fig. 3, lane 14). The other stem flanking the asymmetric bulge (comprised of bases G5-C8 and G28-C31) appears resistant to photomodification by napthol blue black, thionin and nile blue. Only the 3' half of this stem (G28-C31) can be clearly resolved in the autoradiograph. Eosin Y, riboflavin and safranin O yield similar patterns of modification over this short segment. The stops produced by safranin O and riboflavin are intensified when they coincide with photochemical lesions (compare T59 in lanes 10 and 11).

 Primer extension with RT detects modification at all four bases

It is clear that transcriptional pausing occurs at residues other than G, implying photomodification at all four bases. RNA base assignments in Table 1 reflect a preponderance of purine modification, with slightly more guanine than adenine stop sites. An exception to this trend is observed for the napthol blue black modification, where adenine modifications outnumber those at guanine. Among pyrimidines, fewer stops are observed at cytosine than at uracil. Eosin Y, nile blue chloride and napthol blue black fail to photosensitize any cytosine sites. Since all four residues occur in approximately equal amounts for this RNA substrate (G19A21C19U21), the distribution of lesions at G, A, U and C roughly parallels the relative ease of oxidation of the bases (38,39), with cytosine being the most resistant to electron loss.

Photofootprinting of secondary structure

Examination of the autoradiograph in Figure 3 reveals several regions where there is an absence of transcriptional pausing. One such RT 'footprint' corresponds to the region of RNA spanning nucleotides A50–A59. This segment comprises the 3' loop shown in the secondary structure of Figure 2. None of the dyes tested produces any RT detectable lesions over this segment. The stem region flanking this loop, however, is modified to varying extents by all photoreactive dyes, evidenced by pausing at positions corresponding to G60–A65 and G43–C49. The C49-G48 dinucleotide immediately to the 5' side of this loop is only modified by thionin and riboflavin. The 5' loop corresponding to nucleotides G15–A22 likewise appears unreactive towards photomodification. Similarly, we find oxidation occurring in the stem flanking the 3' side of this loop, corresponding to bases G33–C25 of the three base pair stem. All photoreactive dyes appear to modify this stem to some extent. The riboflavin cleavage pattern in this region is displaced slightly to the 3' side relative to the corresponding bands observed with other dyes. Napthol blue black only appears to modify G23 of this stem (Fig. 3, lane 14). The other stem flanking the asymmetric bulge (comprised of bases G5-C8 and G28-C31) appears resistant to photomodification by napthol blue black, thionin and nile blue. Only the 3' half of this stem (G28-C31) can be clearly resolved in the autoradiograph. Eosin Y, riboflavin and safranin O yield similar patterns of modification over this short segment. The stops produced by safranin O and eosin Y are strikingly similar over the entire length of the pGEM(1-80)RNA fragment. Stems appear to be preferential sites of modification in all cases, while RNA loops and bulges do not appear to be targets for the photoreactive dyes reported here.

The region of RNA which demonstrates the greatest variability of photomodification corresponds to the region depicted as single stranded in Figure 2, spanning bases A22–A42. A complete analysis of all possible secondary structures within 15% of the global minimum (p-optimal) was carried out on an SGI Indigo XZ 4400 using MFOLD-2.2 (40). The 3' stem loop was predicted for 12/13 structures found, while the 5' bulged stem loop was predicted for 11/13 structures found.
found in 10/13 structures. The calculated energy of the RNA in Figure 2 is -14.3 kcal/mol. Extremes of reactivity can be seen for thionin (Fig. 3, lane 10) which reacts with almost all the nucleotides in this segment, and nile blue, which is unreactive over this same region (lane 8). The most frequently modified residues in this region are G35C36, and all dyes which react with this segment of the RNA react at either one or both of these positions with the exception of napthol blue black. In this case, modification is seen to occur only at A33 and A39. The varying patterns found for different dyes over this region are difficult to interpret in terms of some obvious structural feature. The order of apparent double helical specificity determined by photofootprinting can be given as nile blue > riboflavin > eosin Y, safranin O > napthol blue black > thionin.

Fluorescence titration studies

Binding to poly(A)poly(U) or poly(G)poly(C) by photoreactive dyes was investigated by fluorescence spectroscopy. Changes in the relative intensity of dye fluorescence upon addition of polynucleotides of varying GC content can be used to identify general features of sequence binding preferences (41,42). Napthol blue black and thionin are not appreciably fluorescent, whereas eosin Y, riboflavin, nile blue and safranin O are strongly fluorescent and thus amenable to this approach. Some modest fluorescent enhancement (~5%) upon the addition of poly(G)poly(C) was observed with riboflavin, while no change in emission intensity could be observed in the presence of poly(A)poly(U). No change in relative emission intensity could be detected for eosin Y and safranin O even upon the addition of >1000-fold excess of homopolymer. This is in contrast to the results obtained with nile blue chloride, shown in Figure 5, where addition of poly(G)poly(C) leads to a >2-fold increase in emission intensity. This is seen to be greater than the observed emission enhancement resulting from addition of an equivalent concentration of poly(A)poly(U). These results demonstrate an enhanced binding affinity for GC rich sites, which are consistent with the photofootprints obtained for this dye.

Scatchard analysis of thionin binding to RNA

The binding of thionin to pGEM(1-80) was investigated by equilibrium dialysis, since the apparent lack of pronounced double stranded specificity could in part be explained by a strong RNA binding affinity. Scatchard plotting of the equilibrium dialysis data (Fig. 4) reveals relatively strong thionin binding with a Ka ~ 6 x 10^-6. Exact determination of the number of binding sites (r intercept) is not possible due to an incomplete isotherm; however, extrapolation from the determined curve indicates multiple binding.

DISCUSSION

Although oxidation reactions of DNA substrates have been well studied (43,44), there are only a few published examples of analogous reactions with RNA. Recently, 8-hydroxyguanosine has been detected in RNA following photooxidation with either methylene blue or rose bengal (10). 5-hydroxyuridine and 5-hydroxycytosine, as well as 8-hydroxyadenine and 8-hydroxyguanine have been identified as 'natural' products in the RNA of Torula yeast (11), although it is possible that their occurrence represents oxidative base damage. It remains to be seen whether other oxidative base modifications detected in DNA (1) will also be found in RNA.

Whatever the precise nature of the photoproducts, it is clear that AMV reverse transcriptase mediated cDNA synthesis can be interrupted at all four nucleotides following oxidative damage. However, it is difficult to determine the relative efficiencies of lesion formation at all four nucleotides solely from the autoradiographic data. This is because the band intensities, which represent the extent of pausing, might not only reflect the relative abundance of a lesion, but may also reflect the relative kinetic ability of reverse transcriptase to read through the different oxidized bases in the template RNA. This would be expected to depend on a number of factors including the structure of the base lesion and its pairing potential with incoming dNTPs. Chemical identification of the
lesions produced and their stoichiometries would help clarify this point and such investigations are in progress.

The variations we observe in the RNA photofootprints would seem to argue against the primary involvement of a diffusible agent, such as singlet oxygen, as the reactive species involved in base modification. If the dyes act primarily as sensitizers for the production of $^{1}O_2$, one would expect the patterns of modification to be identical for all photoreactive dyes and to involve guanosine almost exclusively as the modified base (singlet oxygen has been shown to react almost exclusively with guanines, cf. 45). However, the variations in photofootprints suggest that the dyes are binding to the RNA prior to the photoreaction and that different binding interactions arising from sequence dependent and structural variations dictate the observed specificity. Although we cannot rule out the partial involvement of diffusible oxygen species in every case, the photofootprints obtained for the various dyes support a predominantly Type I reaction pathway, involving direct one electron oxidation of the bases by bound, photoexcited dye triplets. The Scatchard analysis presented for thionin and the fluorescence enhancement observed upon the addition of RNA homopolymers to solutions of nile blue and riboflavin are consistent with a mechanism requiring dye binding prior to RNA base oxidation.

Transcriptase stop sites for most dyes are predominantly at or adjacent to guanosines. This is consistent with a mechanism involving initial binding to either the 3' or 5' side of G, followed by photooxidation of residues adjointing the binding site. This would explain why most stops seem to occur in groups of at least two. Binding adjacent to guanosine is not universal for all photoreactive dyes tested; exceptions were observed for thionin and napthol blue black. In the former case, lesions were detected at A$_{32}$A$_{33}$ and U$_{38}$A$_{39}$, sites which are flanked by non-G nucleotides. These sites are also modified to a lesser extent by napthol blue black.

The specificity for modification at bases within helices is most simply explained by prior association of the dye with the substrate RNA, possibly by intercalation or groove binding. This suggests that the more selective dyes such as nile blue chloride and riboflavin might have utility in the secondary structure mapping of RNA. The origin of the double helical GC binding preference for nile blue and riboflavin indicated by fluorescence studies may be due to a combination of electronic and steric factors which favor interactions at these sites in RNA. GC binding specificity for DNA ligands has been shown to increase as the $\lambda_{\text{max}}$ of the ligand increases. This is because chromophore polarizability increases as a function of wavelength and highly polarizable chromophores can interact more favorably with the more polar GC pairs (46). Examination of the absorbance spectra in Figure 1 reveals that nile blue possesses the most red-shifted maxima of all the photoreactive dyes (640 nm), so it is possible that chromophore polarizability may influence ligand binding to RNA as well. However, riboflavin, characterized by a relatively blue shifted $\lambda_{\text{max}}$, also demonstrates GC specificity. The cloning and sequencing of RNA molecules (‘aptamers’) selected by in vitro evolution methods for riboflavin binding ability from a random pool of RNA oligomers reveals a common G-rich motif (47), supporting our observations. This type of binding specificity has not been observed for solutions of riboflavin with DNA. It is possible that riboflavin binding to RNA is influenced to a greater extent by factors other than chromophore polarizability, for example ribitol side chain interactions in the deep, narrow groove, and thus may not follow the general correlation observed for DNA ligands.

The single stranded loop regions of the pGEM(1-80) RNA are not photooxidized by any of the dyes examined. We interpret this result in terms of the double helical binding preferences of the chromophores. The segment of pGEM(1-80) RNA spanning the 5' and 3' stem loops (A$_{32}$-A$_{42}$) is depicted as single stranded, yet is subject to variable photomodification. Why some dyes are able to photomodify this single stranded RNA segment yet remain unreactive toward the loop regions may at first appear puzzling. One possibility is that the A$_{32}$-A$_{42}$ segment is involved in tertiary interactions, and varying patterns of photomodification reflect differential dye binding affinities for the tertiary structure. Inspection of the sequence A$_{33}$-U$_{38}$ suggests another possibility. This hexameric sequence is self-complementary and pGEM(1-80) dimerization at this sequence would provide an additional duplex binding site. In this case, the propensity of the various dyes to dimerize the pGEM RNA might correlate with the extent of photomodification found in this segment.

Although various experimental and theoretical studies of small molecule/DNA binding have appeared, a complementary understanding of the corresponding complexation phenomena with RNA substrates is lacking. Recent studies have made it clear that the observed DNA binding modes for a particular ligand may not describe the corresponding interactions with RNA. For example, although DAPI binds to the minor groove of poly(dA)poly(dT), it has been found to bind poly(A)poly(U) sequences by intercalation (48). Wilson and coworkers have discussed this problem explicitly in relation to the de novo design of ligands specific for RNA binding (49). The RT assay described here should be of utility in studies of RNA ligand binding, since photofootprinting information can be readily obtained, and since it is a photochemical method, kinetic resolution of high $K_d$ binding sites is possible. Such sites are normally difficult to detect using equilibrium methods because of low occupancy.

Several previous studies of nucleic acid binding molecules have exploited photochemistry in order to obtain information on the resulting complexes. Psoralen derivatives have been used to probe the secondary structure of the Rous sarcoma virus RNA by photocrosslinking (50). Argon laser irradiation of ethidium bromide bound to DNA (51) or small RNA hairpins (52) reveals the binding site location as a phosphate backbone cleavage adjacent to the intercalation site. A transition metal complex of rhodium [bis(phenanthroline)(9,10-phenanthrenequinone diimine)]rhodium(III)) has been shown to yield RNA chain breaks adjacent to bases involved in tertiary interactions upon irradiation with 365 nm light (53). In these cases, however, a mechanism involving preferential oxidation of the sugar rather than the bases is implied since chain breaks are observed. Consistent with the notion that the ribose moiety is the preferred site of oxidative attack in these cases is the detection of released unmodified bases following photolysis of the complexes (50,52,54). Footprinting methods for RNA that do not result in or require chain breakage, such as those described here, might possess some advantage in that sequence-specific binding data can be obtained under mild conditions on the intact molecule. The absence of RNA chain breaks using the conditions we describe is consistent with oxidation of the bases as the principal photochemical event.
CONCLUSIONS AND SUMMARY

Dye photosensitization methods that have been successfully used to generate oxidative base damage in DNA were used to create analogous lesions in a model RNA substrate. Similar to what is observed upon DNA photosensitization, chain breaks are not produced. Primer extension on photolyzed RNA substrates using AMV reverse transcriptase provides a useful assay for the detection of modified bases and our results represent the first report of RT sensitivity to this lesion class. Photomodification produced by dye photosensitization appears to reflect the ease of oxidation of the bases and the binding interactions of the dyes. The specificity reported for nile blue chloride suggests its use as a phototrode for detection of double helical regions in RNA.

Although a variety of oxidizing conditions have been shown to mutate RNA viruses (55,56), severely impair aminoacylation of oxidized tRNAs (57,58), and confer resistance to ribonucleases (59), the biological consequences of RNA oxidative base damage have attracted little attention. The assay described here should make it possible to survey the extent of such damage in cellular RNA in order to assess its relative biological importance. Beyond the potential to interfere with various aspects of translation and RNA processing events, lies the possibility that oxidative base modifications in RNA may contribute to the mutation rate of retroviral genomes. Stable mispairing events analogous to those described for translesional DNA replication can create miscoding errors during RNA template dependent cDNA synthesis.

ACKNOWLEDGEMENTS

We would like to thank Drs A. Bailis, K. Graham and M. R. Lewis for helpful comments. This work was supported in part by the American Lung Association (RG-010-N), The Tobacco Related Disease Research Program (TR2006) of the state of California and the NCI (CA33572). JT is a Young Investigator of the American Lung Association of California.

REFERENCES