Attachment of reporter groups to specific, selected cytidine residues in RNA using a bisulfite-catalyzed transamination reaction

David E. Draper

Johns Hopkins University, Baltimore, MD 21218, USA

Received 18 October 1983; Revised and Accepted 13 December 1983

ABSTRACT

Bisulfite catalyzes transamination of cytidine at the N\textsuperscript{4} position; the suitability of this reaction for attaching reporter groups to selected cytidine residues in RNA molecules has been investigated. Poly(C) is nearly quantitatively converted to the poly (N\textsuperscript{4} aminoethyl-C) derivative after 3 hrs at 42°C with ethylene diamine (pK\textsubscript{a} = 7.6) and bisulfite. This derivative reacts quantitatively with N-hydroxysuccinimide esters; the linkage of a fluorescent dye, nitrobenzofurazan, to cytidine by this reaction is demonstrated. To direct the bisulfite reaction to selected cytidines within a large RNA molecule, the RNA is hybridized to complementary DNA containing a deletion. Only the cytidines in the single strand RNA loop (corresponding to the DNA deletion) are reactive. Two cytidines in the middle of a 340 base RNA fragment from 16S ribosomal RNA have been modified by this technique.

INTRODUCTION

The idea of 'reporter groups' has been a powerful one for analysis of macromolecular conformations and interactions.\textsuperscript{1} The reporter is a spectroscopic probe sensitive to its environment; by covalently linking the reporter to a strategic location in a macromolecule a limited region is monitored as ligands bind or as the macromolecule folds. This technique has been extensively applied to proteins, which frequently have only one or a few reactive sulfhydryls or amines easily linked to a reporter. RNA molecules have not been as tractable. Modification reactions have been confined to the 3' terminus of an RNA (the periodate sensitive 3' ribose) or to uniquely reactive modified bases in some tRNAs (e.g. 4-thiouridine at position 8 of many tRNAs).\textsuperscript{2,3} In RNAs other than tRNA almost no suitably reactive bases occur, so that reporter groups have been placed only at the ends of interesting molecules such as the ribosomal RNAs.\textsuperscript{4} A site-specific method for attaching reporter groups onto one of the four common bases would allow much more detailed physical analysis of RNA structures and interactions than presently possible.

Any base modification method must meet rather demanding criteria to be...
generally useful. The modification reaction must be mild enough to leave large RNA molecules intact; the modified nucleotide must be able to participate in normal base pairing and base stacking interactions; and there must be some way of directing the reaction to a small number of selected bases within a large RNA molecule.

Modification at the N^4 position of cytidine in a bisulfite catalyzed reaction has previously been used to attach reporter groups to RNA. For instance, synthetic polynucleotides containing a fluorescent dye linked to cytidine by this reaction base pair normally and are useful for probing the mRNA binding site of ribosomes. Affinity-labeling reagents and fluorescent dyes have been attached to cytidine residues in tRNA. Unfortunately none of these reaction schemes is well-suited for placement of reporter groups in natural RNAs: either the reaction is very slow and quantitative cytidine modification is difficult to achieve, or the pyrimidine-bisulfite adduct is formed irreversibly. Some amine nucleophiles have been reported to react quantitatively with cytidine with reversible loss of bisulfite, though these have not been useful for further reaction with reporters. In this paper I report a systematic study of the bisulfite-catalyzed transamination reaction, and find that with appropriate reagents and conditions cytidine can be nearly quantitatively attached to reporter groups; the scheme is general for a variety of reporters.

Obtaining site-specific modification of cytidines in RNA molecules is also a problem. Several cytidines in tRNA react readily with bisulfite, and presumably a larger number would react in a larger RNA. In this paper I also show that the specificity of bisulfite for single-stranded bases can be used to direct the reaction to specific regions of an RNA. As an example, two cytidines near the middle of a 340 nucleotide fragment of 16S rRNA have been selectively attached to a fluorescent dye.

**MATERIALS and METHODS**

**Chemicals**

Water was either deionized and distilled from permanganate, or run through a Millipore Milli Q apparatus. Poly(C) was purchased from P-L Biochemicals, nucleosides and nucleotides from Sigma. Sodium bisulfite was Baker reagent grade. Nitrobenzofurazan chloride (NBF-Cl) and chlorotrimethylsilane were from Aldrich. Organic solvents were reagent grade.

**Transamination reaction mixtures**

Rather high concentrations of bisulfite and amines have been used, to
obtain the most rapid reactions possible. These solutions have a significantly lower molar concentration of water than the more dilute solutions used by other workers in studying bisulfite reactions. To easily compare reaction rates, it has been convenient to prepare sets of solutions that have approximately equal total amounts of water, rather than equal volume. For example, 2.37 gm Na₂S₂O₅ and 5.62 gm glycine were dissolved in 20 ml of NaOH solution; varying the NaOH concentration from 0.05 to 1.0 M varied the final pH of the solution between about pH 5.0 and 6.5. This gave a solution about 1.2 molal in bisulfite and 3.7 molal in glycine. The most concentrated ethylene diamine solutions were made by adding, slowly on ice, 8.0 ml concentrated HCl to 7.0 ml water, and then 4.0 ml ethylene diamine (distilled before use). Then 3.79 gm of Na₂S₂O₅ were dissolved and the pH adjusted with a total of 1.0 ml dilute HCl and water. This solution is about 2.5 molal in bisulfite and 3.7 molal in diamine. To all bisulfite solutions was added 1 mg/ml of hydroquinone (dissolved in a small volume of absolute ethanol) to scavenge free radicals. Bisulfite solutions were always made up fresh on the day of use. Reactions were initiated by adding nine volumes of bisulfite-amine solution to one volume of nucleic acid.

Quantitation of reactions

After transamination, reaction mixes were adjusted to pH 8.2 with KOH and allowed to sit 30 min on ice. Nucleo-side and -tide reactions were spotted directly on 500 micron cellulose TLC plates and eluted with 80% isopropanol - 20% water for nucleosides, or 65% isopropanol-2N HCl for nucleotides. Spots corresponding to cytosine, uracil, and product nucleosides or nucleotides were scraped off the glass plate and extracted with 1.5 ml 50 mM sodium citrate, pH 3.0, except for aminoalkyl-CMP products which were extracted with 50 mM sodium borate. The absorbances of the solutions were read at 260 and 280 nm in a Cary 19 spectrophotometer. Occasionally C and U did not completely separate on the TLC plate; two wavelength analysis still allowed their concentrations to be determined. The extinctions of the transamination products were assumed to be the same as the parent cytidine compounds.

Polymers were precipitated from the reaction mix with ethanol (more concentrated bisulfite solutions required dilution with water first), and re-precipitated two times. Complete reversal of the bisulfite adduct was confirmed by the UV absorbance spectrum; the bisulfite-pyrimidine adduct has no peak in the 260 - 270 nm region of the spectrum. To determine base composition, polymers were hydrolyzed with 1 N HCl (2 hr 42°C) and run on TLC.
plates as above; if necessary nucleosides were made from nucleotides by bacterial alkaline phosphatase digestion.

Instead of TLC, some of the experiments used the Pharmacia 'FPLC' apparatus equipped with the Mono Q strongly basic ion exchange column for nucleotide analysis. A 12 ml gradient of 0 to 0.3 M NaCl in 20 mM diethanolamine, pH 8.9, was used. The relative amounts of eluted nucleotides were determined with a Pharmacia UV-1 flow absorbance monitor and Frac-100 peak integrator. The same column was used for analysis of T1 oligomers, using 10 mM Tris pH 7.6 and the NaCl gradient indicated in Figure 7.

Analysis of kinetic data

Rate equations for the set of reactions of the form C+P+U, C+U have been worked out. The three rate constants in this case are k_t (transamination), k_d (deamination of cytidine), and k_c (deamination of product). Semilog plots of the disappearance of cytidine with time gives the sum of deamination and transamination rate constants. A two variable least squares curve fitting routine was used to fit k_t and k_c (with k_t + k_d as a constant) to the plot of product concentration vs. time.

Preparation of fluorescent dye derivatives

To prepare N-hydroxysuccinimide derivatives of the fluorescent dye nitrobenzofurazan (NBF), 0.4 gm NBF-C1 was reacted with 10 mmol of either 6-amino-caproic acid or 5-alanine which had been stirring with 2.5 ml chlorotrimethylsilane in 40 ml dry ethyl acetate. 3 ml triethylamine initiated the reaction which, after workup, gave the expected carboxy-alkyl-NBF derivative. This material, dissolved in dioxane, was reacted with a slight excess of N-hydroxysuccinimide and dicyclohexylcarbodiimide, and the ester product isolated approximately as described for other similar esters. It is stable for several months at 4°C in dimethylformamide.

To react polynucleotides containing aminoalkyl-C derivatives with the esters, one volume of polynucleotide was mixed with an equal volume of 0.2 M Hepes buffer, pH 7.8, two volumes of dimethylsulfoxide, and 0.1 volumes of concentrated ester in dimethylformamide. After reaction for the desired time at room temperature, two times the total reaction volume of absolute ethanol was added, the mixture allowed to sit on ice for 30 min, and the precipitate spun down. Remaining unreacted dye was removed by several ethanol precipitations.

Construction of hybrid

The phage F61 is the fd cloning vector fd106 with a 340 base pair insert, derived from the rrn8 gene for E. coli 16S ribosomal RNA, in the kanR plasmid vector.
gene. The insert covers bases 528 to 867 in the 16S sequence. Fragments of 16S ribosomal RNA complementary to the cloned DNA were made by hybridization of the RNA and DNA, digestion with ribonucleases, and isolation of the RNA fragment by gel filtration in 80% formamide. This procedure will be described in more detail elsewhere (Draper and Kean, ms. in preparation). A 22 base deletion was made in this clone by cutting at the Hind III and Eco RI sites within the insert (positions 652 and 674 in the 16S sequence) filling in the sticky ends with DNA polymerase, and ligating the blunt ends together.

RESULTS AND DISCUSSION

General approach

The two modification schemes considered are shown in Figure 1. In the first step bisulfite catalyzes transamination at the N\(^4\) position of cytidine. Either an amino acid or a diaminoalkane is used to (in effect) attach a reactive group (acid or amine) via a short chain to the N\(^4\) amine. Bisulfite may also catalyze the deamination of cytidine to give uridine; this is a side reaction which must be minimized to obtain a good yield of the modified cytidine. Not shown in this scheme is an intermediate stage in which bisulfite is added across the pyrimidine 5,6 double bond; at mildly basic pH (8.2 - 8.5) the bisulfite is removed and the double bond restored. The second step is the attachment of the probe molecule using either an N-hydroxysuccinimide ester or a water soluble carbodiimide reagent; the net result in either case is the attachment of a probe to the N\(^4\) of cytidine via...
Figure 2. pH dependence of the reaction of cytidine with glycine (3.75 molal) in the presence of bisulfite (1.25 molal), at 42°C. A, time courses of the reactions. •—•, pH 5.21; ▲—▲, pH 5.75; •—•, pH 6.37; ■■■, pH 6.91; ▼▼▼, pH 7.75. R, pH dependence of rate constants. ■—■, transamination rate (kt); •—•, deamination of cytidine (kd); •—•, deamination of carboxymethylcytidine (kc).

an alkyl chain interrupted at some point by a peptide group. The first objective of this work was to find conditions for quantitative conversion of cytidine to transaminated product.

pH dependence of bisulfite-catalyzed transamination and deamination reactions

The bisulfite adduct of cytidine is prone to nucleophilic displacement of the amine at position 4. Both the equilibrium addition of bisulfite and the deamination reaction have been discussed in considerable detail previously.11,16 The bisulfite adduct reaches equilibrium with bisulfite within a few minutes. The extent of adduct formation is influenced by pH, with both the bisulfite - sulfite (pK = 6.2) and cytidine N3 protonation (pK = 4.2) equilibria favoring the adduct at low pH. The deamination step is subject to general base catalysis by sulfite; higher pH enhances deamination by increasing the sulfite concentration. The overall deamination rate thus has an optimum pH near 5.11,16

The transamination reaction should plausibly proceed by the same mechanism as deamination and thus have the same pH dependence. However, in an additional equilibrium the concentration of nucleophile (RNH2) will increase
Table I
Transamination and deamination reaction rates

<table>
<thead>
<tr>
<th>amine</th>
<th>$pK^a$</th>
<th>pH</th>
<th>[amine]$^b$</th>
<th>[HSO$_3$$^b$]</th>
<th>$k_c^c$</th>
<th>$k_d^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>glycine</td>
<td>9.8</td>
<td>6.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.094</td>
<td>0.047</td>
</tr>
<tr>
<td>β-alanine</td>
<td>10.2</td>
<td>6.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.055</td>
<td>0.048</td>
</tr>
<tr>
<td>γ-aminobutyric acid</td>
<td>10.6</td>
<td>6.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.048</td>
<td>0.047</td>
</tr>
<tr>
<td>glycylglycine</td>
<td>8.2</td>
<td>6.3</td>
<td>2.35</td>
<td>1.25</td>
<td>0.13</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.3</td>
<td>2.35</td>
<td>2.50</td>
<td>0.24</td>
<td>0.032</td>
</tr>
<tr>
<td>reaction with poly(C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diaminopropane</td>
<td>9.0</td>
<td>6.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.036</td>
<td>--</td>
</tr>
<tr>
<td>diaminohexane</td>
<td>7.6</td>
<td>6.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.18</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.3</td>
<td>3.75</td>
<td>1.25</td>
<td>0.72</td>
<td>--</td>
</tr>
</tbody>
</table>

$^a$ For diamines, $pK_1$ is given.
$^b$ Concentrations are in approximate molal units, see Materials & Methods.
$^c$ Rate constants are hr$^{-1}$, for reaction at 42°C.

With pH and thus high pH should enhance the transamination rate relative to deamination. Shapiro & Weisgras$^{17}$ noted that transamination is indeed favored at high pH, though no detailed study was made. To quantitate the transamination pH dependence the reaction of glycine with cytidine was monitored and analyzed as described in Materials & Methods, yielding the cytidine deamination rate, transamination rate, and product (carboxymethyl-cytidine) deamination rate. Time courses of reactions at selected pHs are shown in Figure 2A, and the pH dependence of the rate constants in Figure 2B.

Deamination and transamination both slow considerably at high pH, though transamination does drop off much less rapidly and is the dominant reaction above pH 7.0. The product deamination has the same pH dependence as cytidine deamination, but is about an order of magnitude slower. Thus better than 90% yields of product can be obtained above pH 7.0, though the half time for the reaction is more than 15 hours at 42°C. The maximum deamination rate is at about pH 5.3, about the same as found by Shapiro et al,$^{16}$ and the maximum transamination rate is somewhat higher, about pH 5.9. With diamines the maximum transamination rate is at a lower pH, 5.3 (not shown).
Figure 3. Temperature dependence of transamination and deamination reactions. Glycylglycine (2.35 molal) was reacted with cytidine and bisulfite (2.5 molal) at various temperatures. •—•, transamination rate; ■—■; deamination rate.

The transamination reaction rate should depend on the pK of the nucleophile since that pK affects the concentration of reactive RNH₂ present. Table I shows that, among related compounds, lower amine pK is associated with faster transamination. Table I also shows that both transamination and deamination rates are approximately first order in bisulfite concentration, as expected.

Temperature dependence of the reaction

The temperature dependences of the two reactions are shown in Figure 3. The deamination data (WH = 12 kcal/mol) are in reasonable agreement with measurements by others.¹¹,¹⁶ Transamination has a somewhat higher activation enthalpy (ΔH = 18 kcal/mol) and thus is favored (relative to deamination) by higher temperatures.

Reaction with polynucleotides

Since the intended application of this reaction is modification of intact polynucleotides, I have compared the transamination rates of cytidine and poly(C) (Table II). In polymers transamination is slowed by a factor of three to four, while deamination is somewhat more affected. Base stacking and the high negative charge density of polynucleotides probably both limit the accessibility of polynucleotide bases to bisulfite relative to nucleotides.

An interesting aspect of the bisulfite reaction kinetics with polymers
Table II

Comparison of cytidine and poly(C) reaction rates with glycylglycine

<table>
<thead>
<tr>
<th>pH</th>
<th>$k_t$</th>
<th>$k_d$</th>
<th>$k_t$</th>
<th>$k_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.45</td>
<td>0.26</td>
<td>0.24</td>
<td>0.078</td>
<td>0.064</td>
</tr>
<tr>
<td>6.10</td>
<td>0.31</td>
<td>0.045</td>
<td>0.072</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Reaction is at 42°C with 2.5 molal glycylglycine and 2.5 molal bisulfite. Rate constants are hr$^{-1}$.

is that the disappearance of cytidine is always first order to more than 90% completion of the reaction. If the reaction rate of a residue in a polymer were affected by the identity of neighboring residues, then semi-log plots of the reaction time course should be curved, reflecting positive or negative cooperativity. To a first approximation, then, transamination and deamination with polymer residues occur independently.

Attachment of probe molecules

Consideration of the transamination data presented suggests that the diamine reaction scheme is the better method for linking reporter groups to RNA. Ethylene diamine can be made to react with polynucleotides an order of magnitude faster than any amino acid; poly(C) is more than 95% converted to poly(aminoethyl-C) after 3 hr at 42°C in 3.7 molal diamine, 2.5 molal bisulfite, at pH 5.5. The equivalent reaction with glycylglycine takes about 40 hours, and the linker group obtained is bulkier than the ethylene. The conditions described by Kryzosiak et al. for modifying the carboxyl group of a modified base in tRNA have been used successfully for attaching a fluorescent dye derivative, aminopropyl-NBF, to glycylglycine modified poly(C). Because the diamine route seems superior, this carbodiimide approach has not been studied further. There may be situations where, because of the availability of a primary amine on a reporter group, the carbodiimide reaction will be desirable.

The reaction time course of the N-hydroxysuccinimide ester of a fluorescent dye derivative is shown in Figure 4. The reaction proceeds rapidly and quantitatively. Similar results have been obtained with a variety of N-hydroxysuccinimide esters reacting with the diaminopropane derivative of cytidine.
Reaction of specific cytidines in an RNA fragment

To select specific cytidine residues in an RNA molecule for transamination, the single strand specificity of bisulfite can be used to advantage. The initial adduct of bisulfite with a pyrimidine has not been detectable with double helical polynucleotides;\textsuperscript{10,19} the rate of reaction is probably at least three orders of magnitude less than with single stranded polynucleotides. This discrimination has formed the basis for some site-specific mutagenesis methods: bisulfite-catalyzed deamination can be directed to a specific region of a plasmid by making that region single-stranded.\textsuperscript{20} In the same way, the transamination reaction should be localized to single stranded cytidine residues in a polynucleotide.

To test the feasibility of site-directed transamination, the hybrid molecule shown in Figure 5 was constructed. One strand is an fd DNA cloning vector containing an inserted fragment of the 16S ribosomal RNA gene; in the middle of this gene fragment a 22 base deletion has been made. The complementary strand is a fragment of the 16S RNA covering exactly the cloned gene fragment; it of course does not have a deletion and forms a single stranded loop at the position of the DNA deletion. The position and size of the DNA deletion has been confirmed by restriction analysis of the phage replicative form; that the hybridized RNA forms the expected single-stranded loop is confirmed by the experiment shown in Figure 6. Here 5'-\textsuperscript{32}P end
Figure 5. A hybrid between a DNA phage cloning vector containing a 22 base deletion and a complementary 340 base fragment of 16S ribosomal RNA. Arrows indicate the cytidines expected to be modified by bisulfite. The gel shows the sizes of the intact 32P-end labeled RNA fragment (340 bases, lane 2) and the same RNA hybridized to the DNA shown and digested with T1 ribonuclease (126 bases, lane 1).

Figure 6. HPLC analysis of T1 oligomers from the specifically modified RNA fragment. The hybrid shown in Figure 6 was reacted with bisulfite and ethylene diamine, the RNA purified and reacted with the NBF fluorescent dye. This material (30 pmoles), along with 16S RNA carrier (200 pmoles), was digested with T1 ribonuclease and the oligomers run on an anion exchange column as described in Materials and Methods. 0.5 ml fractions were collected, diluted with 1 ml dioxane, and analyzed for fluorescence (excitation, 467 nm; emission, 540 nm). The hatched area indicates fluorescence intensity.
labeled RNA fragment (340 bases) has been hybridized to the DNA with a deletion, digested with T1 ribonuclease, and run on a denaturing gel. The RNA is shortened to approximately the expected size (126 bases).

The RNA - DNA hybrid shown in Figure 5 was reacted with ethylene diamine and bisulfite under conditions which transaminate >95% of cytidine residues in poly(C), as described above. After separating the RNA and DNA strands by gel filtration in 80% formamide, the RNA fragment was reacted with one of the N-hydroxysuccinimide esters of NBF, and then digested with T1 ribonuclease. HPLC analysis of the RNA oligomers produced by the digestion shows that fluorescence is associated only with the pentamer peak (Figure 6), as predicted by the structure of the hybrid. The same experiment has been repeated with 100 pmoles of hybrid having no DNA deletion; none of the T1 oligomers contained detectable (<0.1 pm) fluorescence.

Stability of RNA during transamination

To see if the conditions required for nearly quantitative transamination of cytidine degrade long RNA molecules significantly, 16S RNA was incubated under the same conditions as used above with the RNA-DNA hybrid, and analyzed on denaturing 5% acrylamide, 98% formamide gels (not shown). 85-90% of the RNA is recovered intact. Inclusion of hydroquinone to scavenge free radicals generated by bisulfite is essential; in its absence only a broad smear of degraded RNA is seen on the gel.

CONCLUSION

A general method has been reported for attaching reporter groups to cytidine residues in RNA. Selection of particular cytidines in an RNA depends on the availability of single-stranded DNA complementary to the entire RNA with deletions covering the targeted cytidine. The M13 and fd single-stranded DNA cloning vectors developed for DNA sequence analysis allow cloning of most genes, and oligonucleotide-primed site directed mutagenesis should render any cytidine susceptible to deletion. For the reporter groups to be useful the RNA must be renatured to its biologically active conformation after hybridization to DNA. The 340 nucleotide fragment discussed in this paper has been renatured to a form specifically recognized by ribosomal protein S8 (which normally interacts with this region in intact 16S rRNA) by warming to 65°C in high salt (Draper and Kean, ms. in preparation). Transfer RNA readily renatures under similar conditions. It seems likely that most RNA molecules will renature after hybridization to DNA.

There have been recent proposals for conformational 'switches' in
Nucleic Acids Research

ribsomal RNAs and for unusual RNA structures formed for intron processing. Spectroscopic probes at strategic sites within these RNAs may help to elucidate structures and conformational changes in these RNAs.

ACKNOWLEDGMENTS

I thank Lori Sadler for constructing the fd clone used in these experiments, and Lloyd Greene for help with the kinetic measurements. This work was supported by National Institutes of Health grant GM-29048 and by BRSG S07 RR07041 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health.

REFERENCES