MOLECULAR AND CELLULAR BIOLOGY, Oct. 1985, p. 2842–2846
0270-7306/85/102842-05$02.00/0

Vol. 5, No. 10

Surprising S1-Resistant Trimolecular Hybrids: Potential Complication in Interpretation of S1 Mapping Analyses

MARGARET A. LOPATA, BARBARA SOLLNER-WEBB, AND DON W. CLEVELAND*

Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 15 May 1985/Accepted 26 June 1985

Although the technique of S1 mapping is a powerful analytical tool for the analysis of RNA, we now report a surprising complication involving a trimolecular hybrid between two RNA species and a single DNA probe molecule which, if unrecognized, can lead to misleading interpretations. We document that such trimolecular hybrids can be efficiently formed under some hybridization conditions and that the probe DNA sequence at the junction of the two RNA molecules can be remarkably stable to digestion with S1. Trimolecular hybrids can arise in any instance whenever a distal region of an end-labeled DNA probe is homologous to a moderately abundant RNA in the sample to be analyzed. This situation presents a serious, potential complication for a variety of S1 analyses, particularly those in which DNA transfection has been utilized to reintroduce in vitro-engineered genes into cultured animal cells.

Since its initial popularization by Berk and Sharp (1), the analysis of RNA by S1 nuclease mapping has become a very widely used technique in molecular biology. Although numerous variations of the precise analytical protocol have been formulated (14, 15), all involve hybridization of the RNA sample to be analyzed with a purified DNA probe sequence. Hybridizations are performed under conditions of probe excess so that all RNA molecules capable of hybridizing to the probe are driven into RNA-DNA heteroduplexes. In the usual application, the hybridization mixtures are then digested with the single-strand-specific nuclease S1 under conditions in which only those segments of the radiolabeled probe DNA which are annealed to the RNA are protected from digestion. These S1-resistant fragments are then resolved by gel electrophoresis. The sizes of the protected fragments serve to define the precise regions of probe DNA which are present in the RNA sample, and the signal intensity permits quantitation of the abundance of the homologous RNA species.

As with most techniques in widespread usage, interpretations of many later experiments have relied heavily upon earlier demonstrations of the validity of the protocol as a whole. In the majority of applications, this assumption of validity is no doubt sound. However, we now report a surprising complication involving a trimolecular hybrid between two RNA species and a single probe molecule which, if unrecognized, can lead to remarkably misleading interpretations. This situation can occur in any instance in which different portions of the probe DNA are homologous to two different RNA species, one of which corresponds to a moderately prominent cellular RNA.

To illustrate this potential problem, consider the experiment displayed in Fig. 1. A hybrid gene (pPolI-CAT; Fig. 1A) consisting of a mouse rRNA promoter segment (10) extending from rDNA residues −168 to +57 was cloned adjacent to the coding sequence of the chloramphenicol acetyltransferase (CAT) gene (4). The resultant plasmid was then transiently introduced into cultured animal cells by a DEAE-dextran transfection protocol (7). Mouse L cells were chosen as the recipient line to conform with the reported species specificity of rRNA transcription factors (5, 6, 8, 11).

* Corresponding author.

For isolation of RNA, cells were solubilized in 4 M guanidine thiocyanate at 24 h posttransfection, the DNA was sheared by passage through a 21-gauge needle, and the RNA was pelleted through 5.7 M cesium chloride (2). Ten micrograms of each RNA sample was hybridized (for 12 h in 0.3 M NaCl–100 mM Tris [pH 7.6]–1 mM EDTA at 65°C) to 0.02 pmol of an appropriate single-strand probe of pPolI-CAT DNA which had been end-labeled at a site within the CAT sequence. The mixtures were then brought to 30 mM sodium acetate (pH 4.4), 0.2 M NaCl, and 5 mM $ZnCl_2$, supplemented with 50 μg of denatured bovine DNA per ml, and digested with 60 U of S1 nuclease (Bethesda Research Laboratories) for 1 h at 22°C. Finally, the S1-resistant probe fragments were visualized by autoradiography after separation by denaturing electrophoresis on a 4% polyacrylamide gel containing 8 M urea.

A protected fragment approximately 207 base long would be expected if transcription of the transfected pPolI-CAT gene initiated at the authentic rRNA start site (Fig. 1B-1). In the actual analysis (Fig. 1C, lane 1), this 207-base protected fragment is indeed observed (as is an additional 175-nucleotide fragment which appears to derive from a fortuitous initiation at ~+32; see below). To further authenticate this finding, we repeated the analysis with different amounts of RNA (not shown). The intensity of the 207-nucleotide signal was found to increase linearly with up to 10-fold-larger amounts of assayed RNA, indicating that the hybridizations were in fact conducted in probe excess. Thus, by the generally accepted criteria, these data could be interpreted to demonstrate that a significant amount of transcription of the hybrid gene originated at position +1 of the rDNA promoter.

As an additional control, however, a second experiment was performed with RNA derived from cells transfected with an analogous plasmid construct (pSV2-CAT; Fig. 1A) in which the simian virus 40 early promoter replaces all of the rDNA sequences present in pPolI-CAT (4). The expected result is diagrammed in Fig. 1B-2. Since the transfected sequence does not contain any rRNA sequences, when the probe derived from pPolI-CAT is used, one would not anticipate the protection of any fragment larger than the 151-base region of CAT sequence present on the probe. Surprisingly, the actual experiment gave a very different

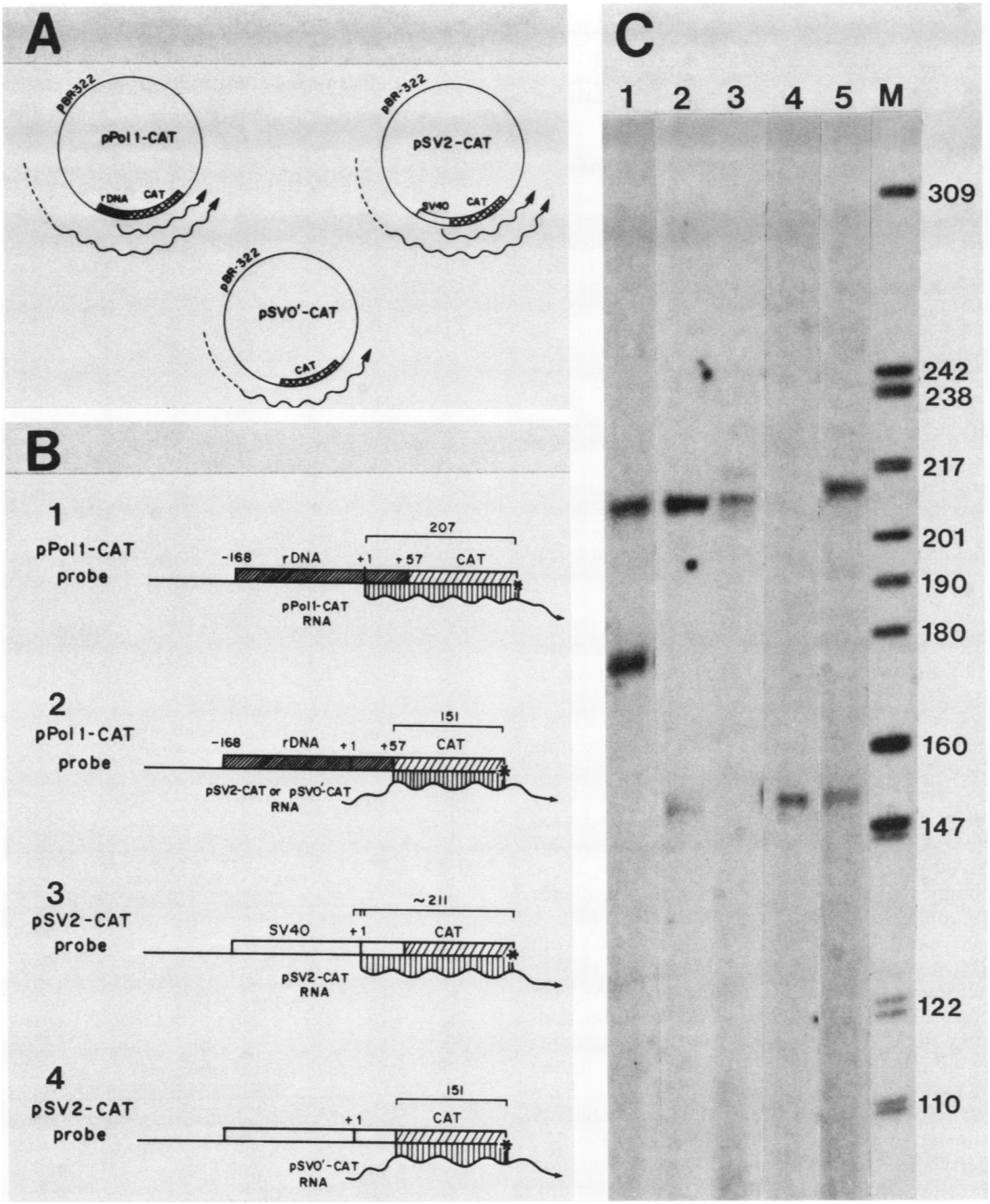

FIG. 1. S1 analysis of RNAs from cells transfected with plasmids pPolI-CAT, pSV2-CAT, or pSV0′-CAT. Mouse L cells were transfected with plasmid DNAs containing the bacterial CAT gene under control of the mouse rRNA promoter sequences (pPolI-CAT) or the simian virus 40 early promoter sequences (pSV2-CAT) or with no eucaryotic promoter (pSV0′-CAT). (A) Schematic drawings of each plasmid and its expected transcripts, including both those starting at known promoters and those reading into this region from possible fortuitous upstream initiations. (B) Expected RNA-DNA hybrids between specific transcripts and probes. All probes were end-labeled at a *Pvu*II site 151 bases within the CAT gene. Panel B-1 shows the 207-base fragment of probe (151 bases of CAT sequence plus an additional 57 bases of the transcribed ribosomal sequence) which is expected when RNA from cells transfected with pPolI-CAT is probed with pPolI-CAT, whereas panel B-3 shows the 151 bases plus an additional 60 to 80 bases of simian virus 40-transcribed sequences which are expected to be protected with probe from pSV2-CAT. (C) Autoradiogram of the S1-resistant fragments of hybrids between: RNA from cells transfected with pPolI-CAT and the pPolI-CAT probe (lane 1, as in B-1); RNA from cells transfected with pSV2-CAT and the pPolI-CAT probe (lane 2, as in B-2); RNA from cells transfected with pSV2-CAT and the probe from pSV2-CAT (lane 3, as in B-3); RNA from cells transfected with pSV0′-CAT and probe from pSV2-CAT (lane 4, as in B-4); RNA from cells transfected with pSV0′-CAT and probe from pPolI-CAT (lane 5, as in B-2).

result (Fig. 1C, lane 2). The majority of the protected probe was again a 207-base fragment, which corresponds in length to protection of the probe up to the precise rRNA transcription initiation site.

A similarly paradoxical situation was obtained when this pPolI-CAT probe was used to analyze RNAs derived from cells transfected with plasmid pSV0′-CAT, which contains the CAT gene and the same pBR322 region as pPolI-CAT

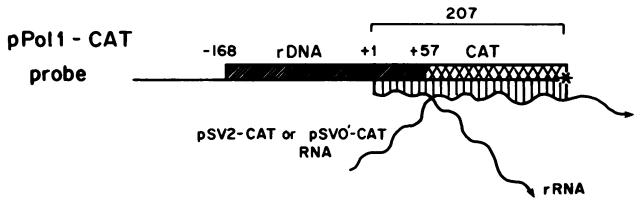

FIG. 2. Structure of the proposed trimolecular hybrid in which a single probe molecule from pPolI-CAT is simultaneously hybridized to two independent RNAs. A schematic drawing of the hybridization of a probe from pPolI-CAT to RNA from mouse L cells transfected with either pSV0′-CAT or pSV2-CAT is displayed. Since the probe sequence contains the first 57 nucleotides of transcribed rDNA sequence and 151 bases of the CAT gene, the region of hybridization between pSV0′-CAT- or pSV2-CAT-directed RNAs and the probe is confined to the 151-base CAT region. However, the 5′ end of the endogenous mouse rRNA primary transcript is moderately abundant in the transfected L-cell RNA, and this RNA can hybridize to the adjoining 57 nucleotides of transcribed rDNA sequences of the probe. The result is a trimolecular hybrid structure which, if not cleaved by S1 nuclease at the discontinuity between the two RNA molecules, will result in a 207-base protected probe fragment corresponding in size to an apparent transcription initiation at the ribosomal transcription start site.

and pSV2-CAT but lacks any authentic eucaryotic promoter element (Fig. 1A). Since this plasmid has pBR322 sequences adjacent to the CAT region, even if RNAs had "fortuitously" initiated in the pBR322 sequences and transcribed into the adjacent CAT sequences, the resultant transcripts would not be expected to protect any more than the 151 residues of CAT sequence homologous to the pPolI-CAT probe (Fig. 1B-2). But again, in addition to the expected 151-base species, a prominent protected fragment of 207 base was observed (Fig. 1C, lane 5). Taken literally, these results would imply that transcription began in a sequence that is not even contained within the template plasmid.

The underlying cause of the unexpected S1 results with RNA from cells transfected either with pSV2-CAT or pSV0′-CAT was further investigated by reanalyzing the RNAs from these transfected cells, but this time with a probe constructed from pSV2-CAT. The expected protected fragments are diagrammed in Fig. 1B-3 and B-4. Reassuringly, with the homologous pSV2-CAT probe, an analysis of RNA from cells transfected with pSV2-CAT yielded the predicted series of closely spaced fragments of 205 to 225 bases (Fig. 1C, lane 3) which derive from the multiple transcription initiation sites known for the simian virus 40 early promoter (3). Moreover, when this pSV2-CAT probe was used to analyze the RNA transcribed from the "promoterless" pSV0′-CAT plasmid (diagrammed in Fig. 1B-4 and shown in Fig. 1C, lane 4) or from the pPolI-CAT plasmid (not shown), only the expected 151-base fragment corresponding to the CAT sequences was protected. Thus, the unexpected S1 protection results appear unique to the pPolI-CAT probe, which carries 57 nucleotides of the transcribed mouse rRNA sequence.

The only plausible explanation for protection of the pPolI-CAT probe extending to the rRNA transcription initiation site when hybridized to RNA from cells transfected with pSV2-CAT or pSV0′-CAT (Fig. 1C, lanes 2 and 5) is that probe DNA hybridized simultaneously to two separate RNA transcripts. The first of these transcripts derives from the transfected gene and protects the probe within the CAT gene sequences, whereas the second, the endogenous rRNA primary transcript, protects the first 57 nucleotides of the rRNA sequences present on the probe. This trimolecular hybrid is diagrammed in Fig. 2. Such trimolecular hybrids can occur because, although the probe is present in vast excess relative to the transcripts from the transfected gene,

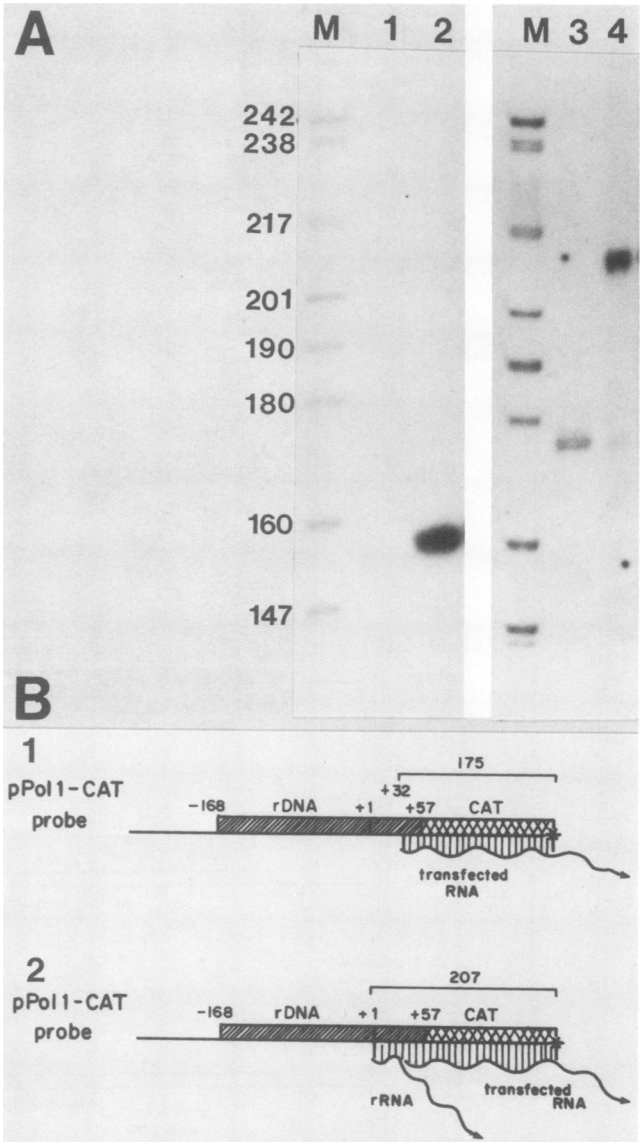

FIG. 3. Demonstration of trimolecular hybrids resistant to S1 nuclease. Mouse L cells were transfected with pPolI-CAT, and at 20 h posttransfection, 20 μM actinomycin D was added to inhibit all further RNA transcription. RNA was then prepared after 1 h more of incubation. That the actinomycin treatment had in fact very markedly depleted the unstable 5′ rDNA transcripts is verified in panel A, lane 1, which shows the S1-resistant fragments obtained from hybridization of this RNA to a 5′ rDNA probe labeled at position +155 of the rRNA transcript. Lane 2, Parallel S1 experiment with RNA from untreated control cells. Lane 3, RNA from the actinomycin D-treated, transfected cells hybridized to the probe from pPolI-CAT. The dominant ~175-base S1-resistant fragment corresponds to a transcription start at position ~+32 within the transcribed rDNA sequence (B-1). Lane 4, S1 experiment in which probe from pPolI-CAT was hybridized to an RNA mixture composed of equal amounts of RNA from the actinomycin D-treated, pPolI-CAT-transfected cells and from untransfected, untreated control cells. The observed 207-base protected fragment arises through a trimolecular hybrid (B-2).

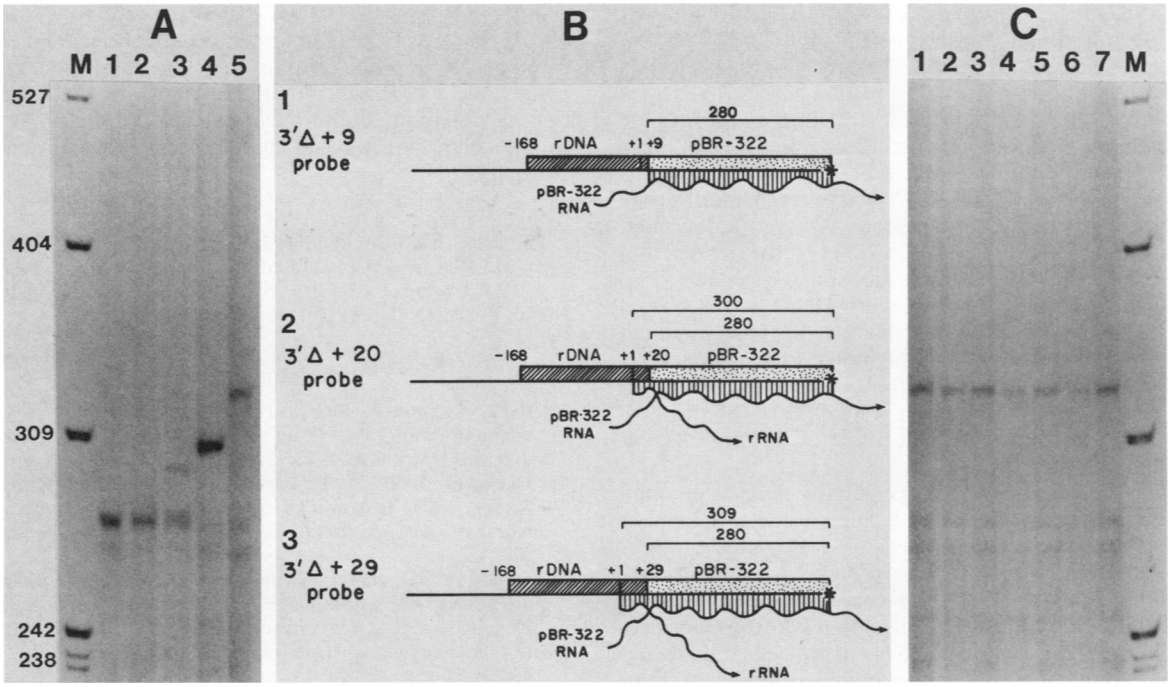

FIG. 4. Determination of the length of RNA sequence sufficient to yield an S1-resistant trimolecular hybrid. Mouse L cells were transfected with pBR322, and the resultant RNA was subjected to S1 analysis with five probes. (A) Results of the S1 analyses with probes carrying 2, 9, 20, 29, or 51 nucleotides of transcribed rDNA sequences (lanes 1 through 5, respectively). If no stable trimolecular structure is formed, only a single 280-base fragment of pBR322 would be protected (B-1). Probes which contain 20 and 29 nucleotides of rRNA sequence form the trimolecular hybrid and yield larger protected fragments and a reduced level of the 280-base band (A, lanes 3 and 4; B-2 and 3, respectively). (C) The stability of the trimolecular hybrids was further investigated by hybridizing the RNA from cells transfected with pBR322 to the probe containing 51 residues of rRNA sequence. The resultant trimolecular hybrids were then subjected to digestion with increasing concentrations of S1 (1× S1 is 60 U/300 μl) or with increasing temperatures of digestion. Lane 1, 1× S1, 24°C; lane 2, 3× S1, 24°C; lane 3, 1× S1, 30°C; lane 4, 3× S1, 30°C; lane 5, 1× S1, 37°C; lane 6, 1× S1, 42°C; lane 7, 0.3× S1, 42°C.

it is not in substantial excess over the 5′ end of the primary rRNA transcript. Thus, although only a small fraction of the probe molecules are hybridized to the transcripts of the CAT gene, most probe molecules are hybridized to the 5′ end of the precursor rRNA. Hence, virtually all probe molecules that are protected by hybridization to the CAT transcript will also be hybridized to the 5′ end of the rRNA and will yield the trimolecular structure.

To demonstrate unambiguously that the cellular rRNA primary transcript is responsible for the unexpected S1 protection results, we performed a reconstruction experiment after elimination of the endogenous rRNA primary transcripts. To this end, a dish of L cells was transfected with pPolI-CAT DNA. At 20 h posttransfection, actinomycin D was added to this dish for the final hour of culturing to inhibit new transcription. RNA was then prepared from this dish and from an untransfected, untreated control dish. These RNAs were first assayed for the presence of endogenous rRNA primary transcripts by S1 analysis with a probe labeled at position +155 of the rRNA primary transcript. As suggested by previous studies (12, 13), treatment with actinomycin D caused complete loss of the endogenous ribosomal primary transcript (Fig. 3A, lane 1), which represents a prevalent RNA in untransfected, untreated control cells (Fig. 3A, lane 2). When analyzed with the pPolI-CAT probe, this same preparation of RNA isolated from the transfected, actinomycin D-treated cells also failed to protect the 207-base segment of probe (corresponding in length to an RNA initiating at the proper rRNA start site); rather, the only protected fragment was ~175 nucleotides and appeared to derive from an initiation at ~+32 of the rDNA sequence in pPolI-CAT (Fig. 3A, lane 3; diagrammed in Fig. 3B-1). Next, an equal amount of the control L-cell RNA was added to this RNA from the transfected, actinomycin D-treated cells, and the mixture was again hybridized with the pPolI-CAT probe and subjected to S1 analysis. In this reconstruction experiment, the protected portion of the probe quantitatively shifted from the position corresponding to ~+32 (Fig. 3A, lane 3) to the position corresponding to +1 of the rRNA sequence (Fig. 3A, lane 4)! The structure of the resultant trimolecular hybrid is diagrammed in Fig. 3B-2.

Thus, we must conclude that (i) trimolecular hybrids are efficiently formed under our hybridization conditions, (ii) S1 nuclease fails to cleave the DNA sequences at the junction of the two RNA molecules, and (iii) the apparent demonstration (Fig. 1C, lane 1) that RNA transcripts directed by pPolI-CAT begin at the authentic rRNA initiation site could be entirely erroneous.

Having determined that such trimolecular hybrids readily form in S1 analyses in which the probe fragment contains only 57 nucleotides of homology to one of the RNAs, we next sought to determine the minimum length of homologous sequence which still yields S1-resistant trimolecular hybrids. To this end, we prepared probes from plasmids containing a systematic series of 3′ rDNA deletions which contain the identical rDNA promoter sequence and between 2 and 51 nucleotides of transcribed rRNA sequence, each cloned into the same site of pBR322. These probes were used to analyze RNA prepared from L cells that had been transfected with

pBR322. Much as the "promoterless" pSV0'-CAT construct had been seen to direct transcripts (Fig. 1C, lanes 4 and 5), we expected that pBR322 would also give rise to transcripts from fortuitous initiation sites. The results of the S1 analyses are shown in Fig. 4A, and schematic drawings of the possible S1-protected fragments are given in Fig. 4B for three of these deletions. When the probe contained only 2 or 9 nucleotides of rRNA sequence, the expected 280-base fragment of probe representing only pBR322 sequences was protected (Fig. 4A, lanes 1 and 2, and Fig. 4B-1). However, for the deletion containing the first 29 bases of the ribosomal transcript (Fig. 4A, lane 4), essentially 100% of the molecules transcribed from pBR322 were driven from a 280- to a 309-base fragment by virtue of simultaneous hybridization to the endogenous rRNA. Even when the probe DNA carried as few as 20 residues of the rRNA sequence, roughly 25% of the protected probe molecules were still found in the trimolecular hybrid (Fig. 4A, lane 3).

Finally, in view of the obvious possibility that the junction region between the two RNA molecules of the trimolecular hybrid might be preferentially sensitive to nuclease digestion, we determined whether the protected fragments observed with trimolecular hybrids could be cleaved by increased extent or stringency of digestion with S1 nuclease. S1 treatment of hybrids between RNA from cells transfected with pBR322 and the probe containing 51 bases of transcribed rRNA sequences (as in Fig. 4A, lane 5) was performed under various conditions. Increasing the amount of S1 by threefold (compare lanes 1 and 2, Fig. 4C) decreased the overall signal, but no preferential loss of the trimolecular species or concomitant protection of a 280-nucleotide fragment was observed. Similarly, increasing the temperature from 24 to 42°C also failed to yield preferential digestion of the trimolecular structure (Fig. 4C).

Thus, our data demonstrate that trimolecular hybrids consisting of two separate RNA species annealed to a single DNA probe molecule can yield DNA fragments which are remarkably resistant to S1 digestion. As in the specific case documented here of mapping the transcripts derived from the pPolI-CAT gene, such trimolecular events can lead to serious errors of interpretation. Moreover, we would hasten to add that this potential complication of S1 analyses is not simply the result of an overwhelming abundance of the rRNA sequence and hence a problem unique to rRNA. Indeed, the 5' end of rRNA is not an extremely abundant RNA. Although mature rRNA makes up over 80% of cellular RNA, molecules containing the extreme 5' end of the rRNA primary transcript are only ~10% of the 45S-rRNA precursor (9), which itself is only ~5 to 10% as prevalent as the mature rRNAs. This calculation suggests that there is only ~0.01 to 0.02 pmol of the 5' end of the rRNA primary transcript per 10 μg of cellular RNA. With this fact in mind, we infer that S1-resistant trimolecular hybrids can also occur whenever a distal region of an end-labeled S1 probe is homologous to other relatively abundant cellular RNAs. These species include tRNAs, 5S RNA, U1 through U6 snRNAs, and small cytoplasmic RNAs such as 7SL RNA. Furthermore, mRNAs are not immune from this complication. A 2,000-base mRNA that makes up only 1 to 2% of the polyadenylated RNA has the same molar abundance in cellular RNA as the rRNA precursor and thus should be similarly prone to forming analogous S1-resistant trimolecular hybrids. We suspect that, as transfection becomes an increasingly popular method for studying initiation, termination, and processing of RNA, such S1-resistant trimolecular hybrids will affect an increasing number of experiments.

This work has been supported by grants GM 27720, GM 34231, and GM 29513 from the National Institutes of Health to B.S.W. and to D.W.C., who is also the recipient of a National Institutes of Health Research Career Development Award.

LITERATURE CITED

1. **Berk, A., and P. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–732.
2. **Chirgwin, J. M., A. E. Przbyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**: 5294–5299.
3. **Ghosh, P. K., and P. Lebowitz.** 1981. Simian virus 40 early mRNA's contain multiple 5' termini upstream and downstream from a Hogness-Goldberg sequence: a shift in 5' termini during the lytic cycle is mediated by large T antigen. J. Virol. **40**:224–240.
4. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
5. **Grummt, I., E. Roth, and M. Paul.** 1982. rRNA transcription in vitro is species specific. Nature (London) **296**:173–174.
6. **Learned, M., and R. Tjian.** 1982. In vitro transcription of human rRNA genes by RNA polymerase. I. J. Mol. Appl. Genet. **1**:575–584.
7. **Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb.** 1984. High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. Nucleic Acids Res. **12**:5707–5717.
8. **Miesfeld, R., and N. Arnheim.** 1984. Species-specific rDNA transcription is due to promoter-specific binding factors. Mol. Cell. Biol. **4**:221–227.
9. **Miller, K., and B. Sollner-Webb.** 1981. Transcription of the mouse rRNA gene by RNA polymerase I. Cell **27**:165–174.
10. **Miller, K. G., J. Tower, and B. Sollner-Webb.** 1985. A complex control region of the mouse rRNA gene directs accurate initiation by RNA polymerase I. Mol. Cell. Biol. **5**:554–562.
11. **Mishima, Y., I. Financsek, R. Kominami, and M. Muramatsu.** 1982. Fractionation and reconstitution of factors required for accurate transcription of mammalian rRNA genes: identification of a species-dependent initiation factor. Nucleic Acids Res. **10**:6659–6670.
12. **Reich, E., and I. Goldberg.** 1964. Actinomycin and nucleic acid function. Prog. Nucleic Acid Res. Mol. Biol. **3**:184–230.
13. **Smale, S. T., and R. Tjian.** 1985. Transcription of herpes simplex virus *tk* sequences under the control of wild-type and mutant human RNA polymerase I promoters. Mol. Cell. Biol. **5**:352–362.
14. **Sollner-Webb, B., and G. Felsenfeld.** 1977. Pancreatic DNase cleavage sites in nuclei. Cell **10**:537–547.
15. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.