MOLECULAR AND CELLULAR BIOLOGY, Aug. 1990, p. 4180–4191
0270-7306/90/084180-12$02.00/0

Vol. 10, No. 8

# N-*myc* mRNA Forms an RNA-RNA Duplex with Endogenous Antisense Transcripts

GEOFFREY W. KRYSTAL,[1]* BARBARA C. ARMSTRONG,[1] AND JAMES F. BATTEY[2]

*Massey Cancer Center, Division of Hematology/Oncology, Medical College of Virginia, and McGuire Veterans Affairs Medical Center, Richmond, Virginia 23249,[1] and Laboratory of Neurochemistry, National Institute of Neurological and Communicative Disorders and Stroke, Bethesda, Maryland 20892[2]*

Received 25 January 1990/Accepted 28 May 1990

**Nuclear runoff transcription studies revealed nearly equivalent sense and antisense transcription across exon 1 of the N-*myc* locus. Antisense primary transcription initiates at multiple sites in intron 1 and gives rise to stable polyadenylated and nonpolyadenylated transcripts. This pattern of antisense transcription, which is directed by RNA polymerase II, is independent of gene amplification and cell type. The nonpolyadenylated antisense transcripts have 5′ ends which are complementary to the 5′ ends of the N-*myc* sense mRNA. We determined, by using an RNase protection technique designed to detect in vivo duplexes, that most of the cytoplasmic nonpolyadenylated antisense RNA exists in an RNA-RNA duplex with approximately 5% of the sense N-*myc* mRNA. Duplex formation appeared to occur with only a subset of the multiple forms of the N-*myc* mRNA, with the precise transcriptional initiation site of the RNA playing a role in determining this selectivity. Cloning of each strand of the RNA-RNA duplex revealed that most duplexes included both exon 1 and intron 1 sequences, suggesting that duplex formation could modulate RNA processing by preserving a population of N-*myc* mRNA which retains intron 1.**

Recently, our understanding of the role of antisense RNA in procaryotic regulatory pathways has greatly expanded. Antisense RNA has been shown to function as a regulator of DNA replication and mRNA translation in a number of experimental systems (18). In addition, expression of a variety of eucaryotic genes has been modulated by injection or transfection of antisense RNA or constructs that express antisense RNA; the mechanisms by which this antisense RNA alters expression of the target gene has been variously attributed to effects on transcription, nuclear processing, nuclear transport, translation, and mRNA stability (18, 41).

Despite the usefulness of artificial antisense constructs in experimental regulation of gene expression, little evidence has been accumulated to suggest a role for naturally occurring antisense RNA in eucaryotic gene regulation. One of the best-characterized examples of an intermolecular interaction of complementary RNA sequences occurs during splicing, when the small nuclear RNAs base pair with complementary sequences of pre-mRNA (28). Another small RNA has been implicated in regulation of translation of the chicken myosin heavy-chain gene (29).

Antisense transcription of the c-*myc* gene, as detected by nuclear runoff experiments, is a well-described phenomenon. Primary antisense transcription 5′ of the major promoters is the most consistent finding, occurring in human HL60 or small-cell lung cancer (SCLC) cells and murine NIH 3T3 cells (5, 25, 32). Attempts to demonstrate a stable RNA derived from primary transcription of an unrearranged allele have been unsuccessful (32). However, studies of rearranged alleles in mouse plasmacytomas have revealed the presence of stable fusion RNAs which initiate on the opposite strand (relative to c-*myc*), within intron 1 of the c-*myc* gene (11, 20). Primary antisense transcription of the 3′ portion of murine c-*myc* has also been demonstrated; this transcription may respond to independent regulatory signals (22, 32).

We have compared and contrasted the regulation of c-*myc*, N-*myc*, and L-*myc* transcription in SCLC cell lines, which, depending on the particular line, can express any of these three *myc* genes (25). One of the differences noted was that while c-*myc* and L-*myc* were, in part, regulated by a block to transcriptional elongation, no such block could be demonstrated during N-*myc* transcription. However, another feature that distinguished between N-*myc* transcription and that of the other *myc* genes was intense antisense transcription over the region of exon 1 and intron 1 of N-*myc* noted in nuclear runoff experiments. It is not surprising, however, that the transcriptional control mechanisms might differ between these genes, which are related in their exon 2 and 3 coding regions but structurally distinct at their 5′ ends. Aside from their lack of sequence homology in exon 1 (13, 19), transcriptional initiation appears to occur differently. Whereas transcriptional initiation occurs at one (L-*myc* [19]) or a few (c-*myc* [4, 6]) discrete sites in c-*myc* and L-*myc* associated with a TATA consensus sequence, N-*myc* transcriptional initiation occurs at more than 20 different sites over a few hundred base pairs of a highly G+C-rich region of DNA unrelated to TATA sequences (23, 38). These start sites may be grouped into two regions separated by approximately 150 nucleotides; initiation in the more 5′ region is associated with RNA processing at a more 5′ splice donor site, and initiation in the 3′ region is associated with RNA processing at a more 3′ splice donor site, giving rise to mature mRNA with different first exons (38). With these unique features of N-*myc* transcription in mind, we decided to determine whether the antisense transcription noted in the in vitro nuclear runoff experiments reflected the presence of antisense transcripts in vivo. Furthermore, we wanted to provide a preliminary structural characterization of the antisense transcripts to determine whether they interact with sense transcripts containing complementary sequences and ultimately to determine the functional significance of these antisense transcripts.

* Corresponding author.

## MATERIALS AND METHODS

**Cell culture.** SCLC cell lines were cultured as previously described (25). Neuroblastoma lines were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum.

**Nuclear runoff transcription.** Nuclear runoff transcription was performed exactly as previously described (25).

**Plasmid construction.** All plasmids used in this study were derived from an N-*myc*-containing genomic bacteriophage clone (CH N-*myc* 5) generated from a human placental Charon 28A library (31). A 2.7-kilobase (kb) *Bgl*II fragment containing the 5'-flanking region and most of exon 1 of the N-*myc* gene was subcloned into pGEM4 (Promega Biotec), as was the adjacent 3' 3.0-kb *Bgl*II fragment containing sequences through the middle of intron 2. A 761-base-pair (bp) *Pvu*II-*Xba*I fragment, including most of exon 1, was excised from the 2.7-kb *Bgl*II plasmid and cloned into pGEM3 to generate the cRNA probes used for the experiments shown in Fig. 3, 6, and 7. A 474-bp *Pvu*II-*Pst*I fragment, including the 5' half of exon 1, was also subcloned into pGEM3 to generate the cRNA probe used in Fig. 9. The 5' portion of the 3.0-kb *Bgl*II fragment, containing mostly intron 1 sequences, was subcloned into pGEM3 by cutting with *Eco*RI to release a fragment between the *Eco*RI site in the pGEM4 polylinker and the *Eco*RI site in intron 1. This plasmid was used to generate the sense probe in Fig. 8; the antisense probe was generated by linearizing the parental plasmid with *Xho*I. The authenticity of all new constructs was verified by comparing restriction maps and sequence information to previously published work (23, 39).

**RNA preparation and analysis.** To obtain maximally intact RNA, cells were lysed in guanidine isothiocyanate and RNA was centrifuged through CsCl by the method of Chirgwin et al. (9) as detailed by Davis et al. (10). Selection of polyadenylated RNA was performed by the method of Aviv and Leder (1). Northern (RNA) blots, nick translation, and production of single-stranded DNA probes from M13 vectors were done as described by Davis et al. (10).

Analytic RNase protection assays were performed as described by Winter et al. (43), with the following exceptions. Radioactively labeled ([$^{32}$P]UTP) cRNA probes were made by in vitro transcription by using SP6 or T7 RNA polymerase (Promega Biotec) as recommended by the manufacturer after linearization of the above-described plasmids at a restriction site in the polylinker (except where noted) distal to the appropriate promoter for making sense- or antisense-specific probes; annealing of the probe ($10^5$ cpm) to the RNA sample was performed for 16 h at 50 to 55°C; RNase digestion was performed with 100 μg of RNase A per ml (RNase $T_1$ was omitted because at this concentration of RNase A it was found not to alter the digestion pattern) in 200 mM NaCl–100 mM LiCl–10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA for 30 min at 30°C.

The double RNase protection procedure for detecting in vivo RNA-RNA duplexes was performed as follows. Cells were lysed with 10 to 15 strokes of a Dounce homogenizer (loose-fitting pestle) in 5 volumes of 10 mM Tris hydrochloride (pH 7.4)–3 mM $CaCl_2$–2 mM $MgCl_2$–0.5% Nonidet P-40 buffer, and the nuclei were pelleted at 500 × *g* for 10 min at 4°C. The supernatant containing the cytoplasmic RNA was made 0.25 M NaCl, 200 μg of RNase A (Sigma Chemical Co.) per ml and 40 U of RQ1 DNase (Promega Biotec) per ml were added, and the mixture was incubated at 37°C for 30 min. The supernatant was then made 200 mM Tris hydrochloride (pH 7.4), 25 mM EDTA, 1% sodium dodecyl sulfate; 0.4 mg of proteinase K per ml was added, and the incubation was continued for an additional 30 min at 42°C. The supernatant was then phenol-chloroform extracted, and the RNase A-resistant double-stranded RNA (and some small precipitable oligonucleotide digestion products) was precipitated with an equal volume of isopropanol. After being washed with 80% ethanol, the double-stranded RNA was dried and then suspended in water and the concentration was determined by the optical density at 260 nm. This RNA was then heated to 90°C for 2 min and used in the analytical RNase protection assay detailed above, with either sense- or antisense-specific radiolabeled cRNA probes. Control RNA was obtained by the same procedure, except that during the initial incubation RNasin RNase inhibitor (1,000 U/ml; Promega Biotec) was added instead of RNase A.

RNase III digestion was accomplished by suspending 30 μg of RNase A-resistant RNA in buffer containing 10 mM Tris hydrochloride (pH 7.9), 10 mM $MgCl_2$, 1 mM β-mercaptoethanol, 5% glycerol, 20 mM KCl, and increasing amounts of RNase III (0.45, 0.90, and 1.8 μg of protein); incubation was for 30 min at 37°C. Controls consisted of untreated RNase A-resistant RNA and resistant RNA retreated with 10 μg of RNase A. After digestion, analytic RNase protection was performed as detailed above.

**Primer extension.** Specific oligonucleotide (250 ng) was 5' end labeled by a kinase reaction (10). Approximately $10^5$ cpm of oligonucleotide was incubated for 16 h with 50 μg of total RNA in 10 μl of 0.4 M NaCl–8 mM PIPES [piperazine-*N*-*N*'-bis(2-ethanesulfonic acid)] buffer (pH 6.7). The incubation temperature was determined by calculating the melting temperature for the oligonucleotide (10) and subtracting 10°C (generally, 55 to 65°C). The annealing mixture was diluted to 100 μl with addition of 1,000 U of Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories, Inc.) and 0.5 mM deoxynucleoside triphosphates with the manufacturer-supplied buffer. The extension reaction was performed at 37°C for 60 min and terminated by heating to 90°C for 2 min. A 10-μg sample of RNase A (DNase free) was added, and the incubation was continued for 15 min at 37°C. The mixture was then made 0.5% sodium dodecyl sulfate, 30 μg of proteinase K was added, and the incubation was continued for 15 min at 42°C. The primer-extended species were then phenol-chloroform extracted, ethanol precipitated, and electrophoresed on a 5% acrylamide–8 M urea sequencing gel (10).

**Cloning and sequencing of RNA duplexes.** Cloning of double-stranded RNA was accomplished by using a modification of the RACE protocol for polymerase chain reaction (PCR) amplification of the 5' ends of cDNAs (17). An oligonucleotide primer (5 ng) specific for either the sense or antisense strand (for sequence details, see Fig. 10) was annealed to 30 μg of RNase A-resistant RNA, and cDNA was synthesized by using the primer extension protocol detailed above. cDNA was purified by annealing the reaction mixture to glass powder (GENECLEAN; Bio 101), washing it as recommended by the manufacturer, and eluting it with TE buffer (10). The cDNA was then tailed with dATP as described by Frohman et al. (17). One-quarter of the tailing reaction was diluted to a 100-μl total volume and used in the PCR reaction. The PCR reaction mixture contained a 1 μM concentration of orientation-specific primer designed to hybridize 3' of the primer used for cDNA synthesis (see Fig. 10); a 1 μM concentration of a $(dT)_{17}$ primer-adapter containing *Hin*dIII, *Sal*I, and *Xho*I cloning sites (5'-GACTC GAGTCGACAAGCTTTTTTTTTTTTTTTT-3'); 50 mM KCl; 10 mM Tris hydrochloride (pH 8.4); 1.5 mM $MgCl_2$;

0.01% gelatin; and 2.5 U of Taq polymerase (Amplitaq; Cetus Corp.). Amplification was accomplished in a programmable thermal reactor by using a step program (95°C, 1.5 min; 53°C, 2 min; 72°C, 2 min) for 30 cycles, with a final extension for 10 min at 72°C. A long denaturation step was absolutely required for efficient amplification, presumably because of the G+C-rich nature of the sequence. Blunt ends on the amplified DNA were assured by incubating the phenol-choloform-extracted PCR products with T4 DNA polymerase as described by Davis et al. (10). The amplified sense cDNA was cut with *Sal*I and *Bam*HI (a site in orientation-specific primer ε; see Fig. 10) restriction endonucleases and ligated into a pGEM3 vector (Promega Biotec) cut with the same enzymes. The antisense cDNA was cut with *Sal*I alone (orientation-specific primer σ) or *Sal*I and *Bam*HI (orientation-specific primer β) and ligated into *Sal*I-*Sma*I-cut pGEM3 or *Sal*I-*Bam*HI-cut pGEM3, respectively. Colonies were screened for inserts by using a *Pvu*II-*Pst*I fragment spanning the 5′ portion of exon 1 (sense cDNA), a *Bam*HI-*Bgl*II fragment spanning the 3′ portion of exon 1 (antisense cDNA derived from primer pair α-β), and a 26-base synthetic oligonucleotide (AATCTGGGTCACGG CTGCTCCAGCTT) representing sequences in the 5′ portion of intron 1 (antisense cDNA derived from primer pair π/σ). Miniprep DNA (10) was purified by using glass powder as detailed above and sequenced by the chain termination technique by using either an SP6 or a T7 sequencing primer (Promega Biotec). Approximately half of the clones were sequenced manually with a Bethesda Research Laboratories sequencing kit and the directions for double-stranded templates. The other clones were sequenced by using a Genesis 2000 automated DNA sequencer (Du Pont Co.) with the manufacturer-supplied kit and directions for double-stranded templates.

## RESULTS

**Antisense transcription, which is mediated by RNA polymerase II, is independent of gene amplification and cell type.** Preliminary nuclear runoff transcription assays using N-*myc*-amplified SCLC cell line H249 had demonstrated extensive antisense transcription 5′ of the *Eco*RI site in intron 1, as well as some antisense transcription between the *Eco*RI site and a *Bam*HI site in exon 2 (see Fig. 4 for a restriction map); there was no detectable antisense transcription of sequences corresponding to the coding region of exon 3 (25; Fig. 1, lane A). To map the region of antisense transcription more precisely, we moved the 5′ border of our exon 2 target from the *Eco*RI site to an *Xho*I site in exon 2 proper (Fig. 1, lane B). When this was done, most of the antisense transcription detected by our exon 2 target disappeared, indicating that most of the antisense transcripts were initiated 5′ of the *Xho*I site in exon 2. Densitometric analysis of lanes A and B revealed, after normalization for differences in exposure using the exon 1 sense signal and correction for uridine content, that antisense transcription was three times greater between the *Eco*RI and *Xho*I sites than between the *Xho*I and *Bam*HI sites. To determine whether the antisense transcription was related to gene amplification, we studied SCLC cell line H187, which expresses N-*myc* without gene amplification (Fig. 1, lane C). Although the absolute amount of primary transcription relative to actin was less in H187 (compare lanes B and C), the pattern of transcription, in particular, the ratio of exon 1 sense to antisense transcription, was identical, indicating that antisense transcription occurs independently of gene amplification. To determine whether this phenomenon was peculiar to SCLC cells, we performed the same experiment with a classic neuroblastoma cell line, Lan5, which is amplified for the N-*myc* gene (lane D). The results were identical to those obtained with the SCLC lines, indicating that antisense transcription is characteristic of the N-*myc* locus and independent of the cell type. Furthermore, both sense and antisense transcriptions are mediated by RNA polymerase II, since they were completely inhibited by 2 μg of α-amanitin per ml, which did not affect polymerase I (28S)- or polymerase III (5S)-mediated transcription (lane E).

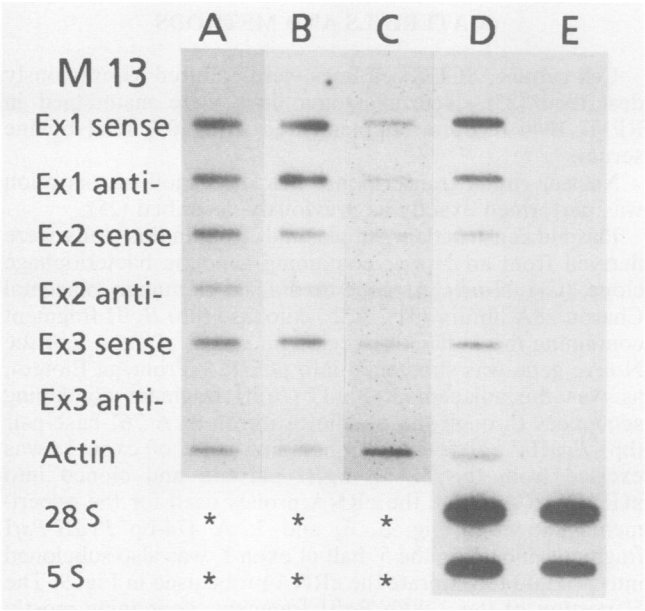


FIG. 1. Nuclear runoff transcription of the N-*myc* gene. Nuclear runoff transcription was performed by using nuclei isolated from H249, an SCLC cell line amplified for the N-*myc* gene (lanes A and B); H187, an SCLC cell line which expresses N-*myc* without gene amplification (lane C); and Lan5, a neuroblastoma cell line amplified for the N-*myc* gene (lane D; lane E, with 2 μg of α-amanitin per ml). All N-*myc* targets are strand-specific M13 clones covering the following sequences (25; see the block diagram in Fig. 4): exon 1, *Bam*HI-*Eco*RI; exon 2 (lane A), *Eco*RI-*Bam*HI; exon 2 (lanes B to E), *Xho*I-*Bam*HI; exon 3, *Bgl*II-*Pst*I. *, No experimental sample.

**Multiple N-*myc* transcripts detected by Northern blot analysis.** Preliminary experiments in our laboratory had detected multiple N-*myc* transcripts (31), in addition to the well-characterized 3.1-kb mRNA. To characterize these transcripts further, we electrophoretically separated 10 μg of cytoplasmic, nuclear, polyadenylated, and nonpolyadenylated RNAs from cell line H249, transferred them to nitrocellulose, and hybridized successive 5′-to-3′ N-*myc* genomic nick-translated probes to the blot (Fig. 2A). These double-stranded probes should detect both sense and antisense transcripts. All of the probes detected the 6.6-kb N-*myc* primary transcript (39) in the nuclear and polyadenylated samples. The *Pvu*II-*Pst*I probe, encompassing the 5′ portion of exon 1, detected primarily the previously described 3.1-kb N-*myc* mRNA (23, 39), which was efficiently transported to the cytoplasm; this was also the predominate mRNA detected by the *Eco*RI-*Bam*HI probe encompassing exon 2 sequences. The *Bam*HI-*Bgl*II probe, encompassing the 3′ portion of exon 1, rather than detecting the 3.1-kb RNA, detected predominantly a 4.2-kb transcript, also strongly detected by the adjacent *Bgl*II-*Eco*RI probe, representing sequences in the 5′ portion of intron 1. This 4.2-kb RNA was

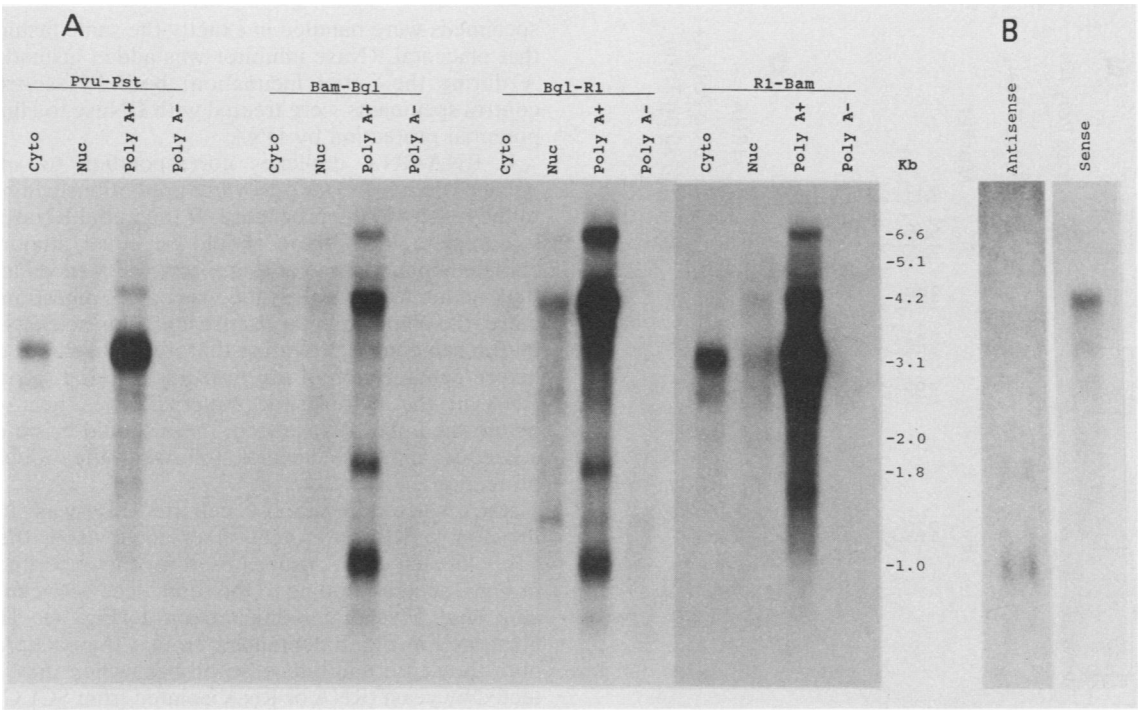


FIG. 2. Northern blot analysis reveals multiple N-*myc* transcripts. (A) Ten-microgram samples of cytoplasmic, nuclear, polyadenylated, and nonpolyadenylated RNAs from cell line H249 were electrophoretically resolved, transferred to nitrocellulose, and hybridized to successive 5′-to-3′ N-*myc* genomic double-stranded probes. The probes represent the following regions: *Pvu*II-*Pst*I, 5′ portion of exon 1; *Bam*HI-*Bgl*II, 3′ portion of exon 1; *Bgl*II-*Eco*RI, 5′ portion of intron 1; *Eco*RI-*Bam*HI, 3′ portion of intron 1 and most of exon 2 (see Fig. 4). (B) Five micrograms of polyadenylated RNA from cell line H249 was resolved and hybridized to antisense- or sense-specific single-stranded probes encompassing sequences between the *Bam*HI site in exon 1 and the *Eco*RI site in intron 1.

not detected by intron 2-specific probes (data not shown), indicating that it represents a species which retains intron 1 but has intron 2 removed; the sizes of the primary transcript (6.6 kb) and intron 2 (2.6 kb; [39]) are consistent with this conclusion. Although much of this RNA appears to be an intermediate in the processing of the primary transcript to its mature form because of its predominantly nuclear localization, some cytoplasmic 4.2-kb RNA was detected, especially with the *Bgl*II-*Eco*RI probe. Also detected with the *Bam*HI-*Bgl*II and *Bgl*II-*Eco*RI probes were 1.8- and 1.0-kb species, which were not detected by the adjacent 5′ and 3′ probes. When polyadenylated RNA was probed with a singled-stranded probe specific for antisense transcripts between the *Bam*HI and *Eco*RI sites, the 1.8- and 1.0-kb bands were detected (Fig. 2B), whereas the 6.6-, 4.2-, and 3.1-kb transcripts were not (but they were detected by a sense-specific probe), indicating that the smaller transcripts were polyadenylated antisense RNA including sequences derived from the 3′ portion of exon 1 and the 5′ portion of intron 1. These polyadenylated antisense transcripts were detected only when relatively large amounts of polyadenylated RNA were analyzed, indicating that they are rare. However, when modest amounts of total or nonpolyadenylated RNA were probed with the same sequences, a heterogeneous population of species was reproducibly detected migrating at around 1 kb; this is best seen in Fig. 2A in the nonpolyadenylated lane probed with the *Bgl*II-*Eco*RI fragment.

**Most of the stable antisense transcripts are not polyadenylated.** To characterize the structure of the antisense transcripts further, we used an RNase protection assay of RNA isolated by the guanidine isothiocyanate method from cell line H249. To interpret the data, however, it was necessary to develop landmarks within the locus by mapping the N-*myc* sense transcripts. Figure 3a illustrates the major sense RNA species detected with a cRNA probe extending from the *Xba*I site at the 3′ border of exon 1 to a *Pvu*II site in the 5′-flanking region of the gene. The group of protected fragments labeled A corresponds to major transcripts which initiate in the 5′ grouping of initiation sites (labeled $I_1$ in Fig. 4) and splice at the 5′ splice donor site (SD1). The protected fragments labeled B in Fig. 3a represent the major transcripts which initiate in the 3′ group of initiation sites ($I_2$) and splice at the 3′ splice donor site (SD2) or retain intron 1. The protected fragments labeled C represent transcripts which initiate in $I_1$ and either splice at SD2 or retain intron 1. These conclusions were supported by extensive mapping of this area with smaller overlapping cRNA probes and are consistent with those previously published (23, 38), except for detection of species C, which represents predominately the primary transcript, and a 4.2-kb form which retains intron 1 but lacks intron 2 (Fig. 2). All of the major species detected with the sense probe are preferentially polyadenylated, as evidenced by the enrichment in the sample containing 5 μg of poly(A)$^+$ as opposed to that containing 30 μg of total RNA. When the same RNA was probed with a cRNA probe designed to detect antisense transcripts (Fig. 3b), multiple protected species were seen, indicating the existence of stable antisense RNA species. In contrast to the sense RNA, however, most of the antisense RNA detected by this probe was nonpolyadenylated, with the exception of a triplet at approximately 80 bases. This triplet represents a component of the 1.0- and 1.8-kb polyadenylated antisense transcripts detected on a Northern blot (Fig. 2). The sizes of the protected bands made it clear that the sense and antisense

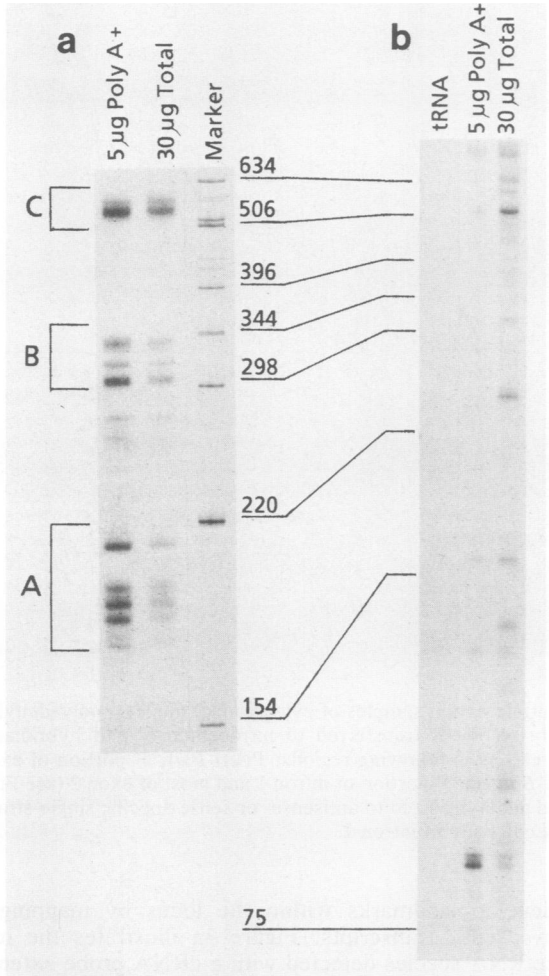


FIG. 3. RNase protection analysis of the 5′ region of the N-*myc* locus. (a) Analytic RNase protection of an 805-base sense-specific cRNA probe (transcribed from the *Xba*I site at the 3′ border of exon 1 to the *Pvu*II site 5′ of the gene; see Fig. 4) by 5 μg of polyadenylated RNA or 30 μg of total RNA isolated from cell line H249. The lettered groupings of bands represent specifically initiated and spliced transcripts, as described in the text and diagrammed in Fig. 4. (b) Assay using the same samples of RNA to protect an antisense-specific probe (transcribed from the *Pvu*II site to the *Xba*I site) encompassing the same sequences. The exposure time of panel a was approximately 1/20 that of panel b. The numbers between the panels indicate molecular sizes in bases.

RNAs contain overlapping complementary sequence regions which could potentially form an RNA-RNA duplex. To test this possibility, we designed a technique which could detect and characterize double-stranded RNA molecules present in vivo.

**N-*myc* mRNA forms a duplex with antisense transcripts in vivo.** We used a double RNase protection technique, illustrated in Fig. 5, to detect and analyze RNA-RNA duplex molecules formed in vivo. Cytoplasmic RNA was crudely isolated without denaturing any potential double-stranded RNA and then subjected to high-concentration RNase A digestion. The RNase-resistant RNA, along with some of the small oligonucleotide products of digestion, was precipitated with isopropanol. The resistant RNA was then denatured and annealed to a uniformly labeled cRNA probe, and a standard RNase protection assay was performed. Control specimens were handled in exactly the same fashion, except that placental RNase inhibitor was added instead of RNase A during the initial incubation; both RNase-treated and control specimens were treated with DNase to eliminate any potential protection by DNA.

If RNA-RNA duplexes corresponding to specific sequences do exist in vivo, several predictions can be made as to the results of the procedure. If the initial RNase digestion is complete, then there should be equal amounts of the complementary sense and antisense sequences after digestion, no matter what the ratio was before digestion. Furthermore, the complementary sense and antisense RNAs should be the same size, providing that all RNase A is eliminated before denaturation of the resistant material. Finally, if the RNA in the cytoplasmic supernatant is heat denatured before the initial RNase step, there should be no detectable sense or antisense species following the double RNase protection.

RNA isolated from SCLC cell line H249 was subjected to the double RNase protection protocol using the radioactively labeled cRNA probes used for the experiment shown in Fig. 3, corresponding to the sequences between the *Pvu*II and *Xba*I sites that bracket exon 1 (Fig. 4). The results illustrated in Fig. 6 demonstrated that the background with both the sense and antisense probes, when they were protected by yeast tRNA or RNA isolated from SCLC line H146 (a c-*myc*-expressing cell line), was negligible. Thirty micrograms of H249 RNA (as determined by the optical density at 260 nm) treated as a control, treated with RNase A (RNased), or heat denatured before RNase treatment (heated) was annealed to each probe. After analytic RNase digestion, the control specimens demonstrated essentially the same pattern that was noted in Fig. 3, with the exception of some mild degradation, noted especially with the sense probe. Note the large excess of sense RNA versus antisense RNA in the control specimens (the probes were of equal sizes and similar specific activities). However, after RNase digestion of the cytoplasmic RNA, both sense and antisense probes appeared to detect nearly equivalent amounts of resistant material, as expected for an RNA-RNA duplex encompassing these sequences. Furthermore, the major species detected with both probes after RNase treatment were the same size (within the ability to size complementary sequences on this gel system), also expected if duplex molecules exist. Minor bands which do not correspond probably represent degradation during incubation with the radioactive probe, since at this point, the RNA has been denatured, rendering it sensitive to residual traces of the high concentration of RNase A used in the initial digestion. Finally, as predicted, heating the cytoplasmic RNA to 90°C for 3 min before RNase treatment abolished all protected species. To determine whether the duplexes were formed after cell lysis, we added an excess of labeled cRNA probe to our cells during lysis. After RNase digestion, there was no protection of either the sense or the antisense probe, indicating that duplexes did not form during RNA preparation (data not shown).

Quantitatively, approximately 5% (3 to 8%) of the sense mRNA appeared to be duplexed with complementary antisense RNA, as determined by densitometry of appropriately exposed autoradiograms. Interestingly, duplex formation appeared to be selective for certain sense mRNA species (implied by a marked reduction in the number of species after RNase treatment, with complete loss of some of the most abundant ones). On the other hand, the antisense species showed virtually no diminution, with the exception

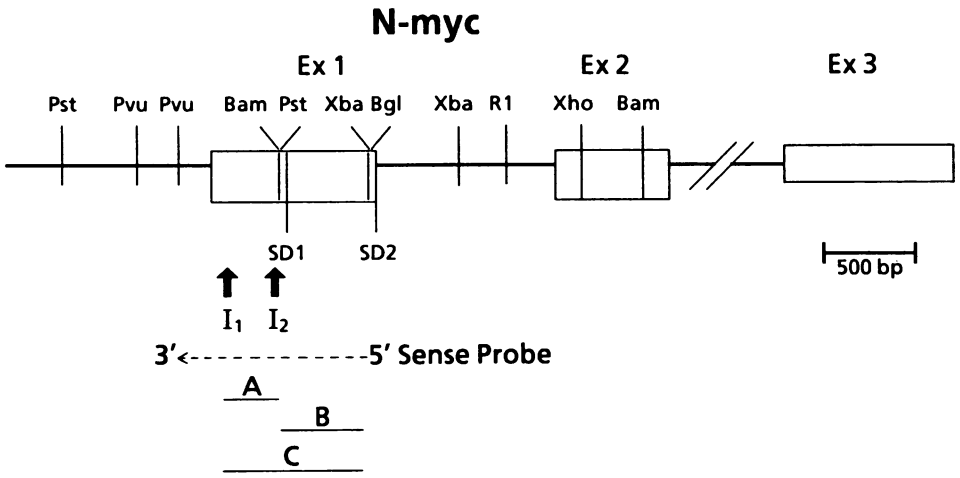


FIG. 4. Restriction map of the 5' portion of the N-*myc* locus. Restriction sites pertinent to this study are illustrated on a block diagram of the N-*myc* locus drawn approximately to scale, except for exon 3. The regions labeled $I_1$ and $I_2$ represent the alternative regions of transcriptional initiation; the points labeled SD1 and SD2 represent the alternative splice donor sites described by Stanton and Bishop (38). The sense-specific probe used in the experiment whose results are shown in Fig. 3 and the expected protected fragments derived from the alternate transcriptional initiation and splice sites are illustrated; the lettering corresponds to the groups of bands in Fig. 3a. Abbreviations: Pst, *Pst*I; Pvu, *Pvu*II; Bam, *Bam*HI; Xba, *Xba*I; Bgl, *Bgl*II; R1, *Eco*RI; Xho, *Xho*I.

of the triplet representing polyadenylated RNA (Fig. 3) at 80 bases. In fact, there was reproducible enrichment of antisense species after RNase treatment, indicating that most of the stable antisense RNA exists in a duplex. The quantity of duplexed RNA, which is resistant to RNase A, should remain relatively constant while the bulk of single-stranded RNA would be digested into species inefficiently precipitated with isopropanol, thus accounting for the enrichment.

Further evidence that the sense and antisense RNAs detected as described above exist as a duplex would be to demonstrate sensitivity to double-strand-specific RNase. We therefore subjected 30-μg samples of purified RNase A-resistant RNA to digestion with a purified preparation of RNase III. Increasing amounts of RNase III caused progressive loss of species detected by an antisense-specific probe (Fig. 7, lanes B through D); equivalent results were obtained with a sense-specific probe (data not shown). Purified RNase A-resistant material also remained resistant upon redigestion with RNase A (Fig. 7, lane E). In addition, purified RNase A-resistant RNA from a c-*myc*-expressing cell line, H146, also gave no protection of an N-*myc*-specific probe (lane F).

**Detection of initiation sites for antisense transcription.** To begin to map the extent of duplex formation, we sought to characterize the major initiation sites for antisense transcription. We knew from the nuclear runoff studies that there was very little antisense primary transcription 3' (sense orientation) of the *Xho*I site in exon 2 (Fig. 1), so we transcribed a labeled cRNA probe from the *Bgl*II site in exon 1 to the *Xho*I site in exon 2 (Fig. 4) to detect the 5' ends of the antisense RNA. The control lane of the antisense panel in Fig. 8 illustrates the results of the RNase protection assay in which this probe was annealed to 30 μg of RNA isolated from line H249. The major initiation site appears to be approximately 420 bp 3' (sense orientation) of the *Bgl*II site, in the proximity of the *Xba*I site in intron 1, with other minor sites between *Bgl*II and *Xba*I. Longer exposures, as well as other experiments using guanidine isothiocyanate-isolated RNA, illustrated minor initiation sites in the vicinity of the *Eco*RI site in the middle of intron 1 as well (data not shown).

To determine whether these antisense species were in a duplex with sense transcripts in this region, we hybridized RNase-digested RNA to this antisense probe. Most antisense species are preserved after RNase digestion, indicating that they are in a duplex, with the exception of a grouping of transcripts at approximately 220 bp. To confirm that the duplex contained sense N-*myc* transcripts, we transcribed a

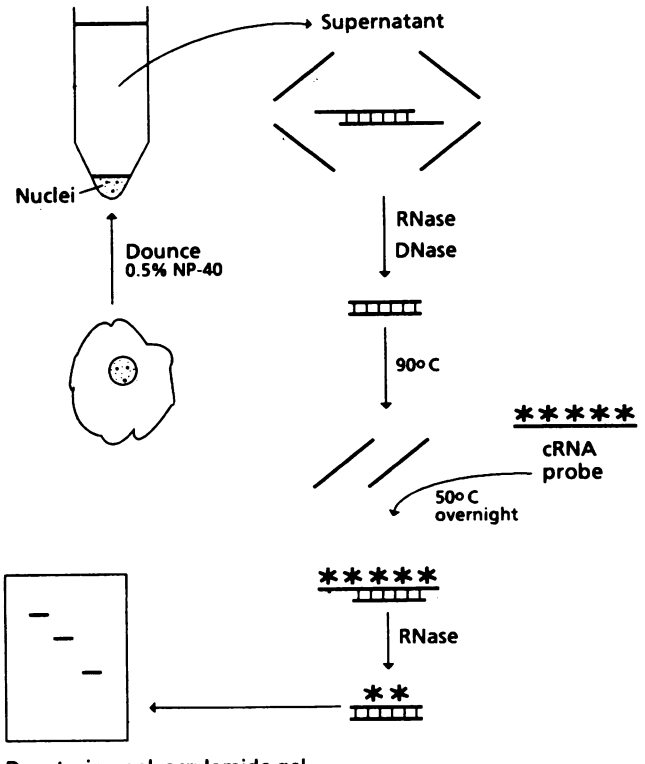


FIG. 5. Double RNase protection protocol for detection of specific in vivo RNA-RNA duplexes. NP-40, Nonidet P-40.

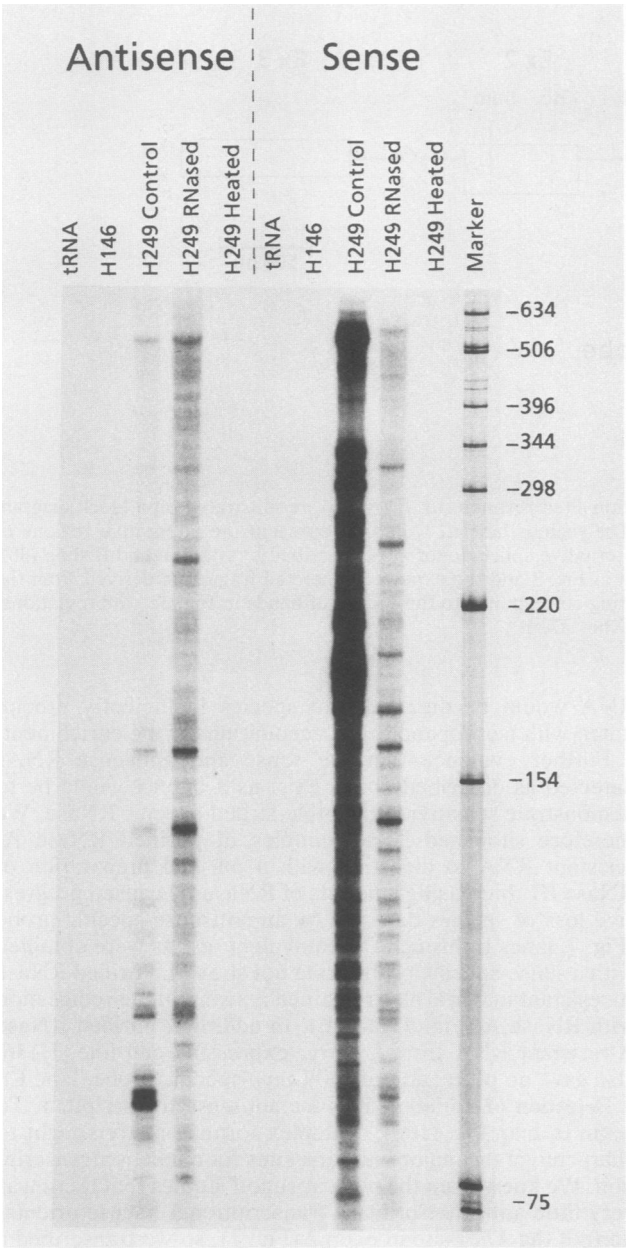


FIG. 6. Detection of in vivo RNA-RNA duplexes corresponding to N-*myc* exon 1 sequences. SCLC cell line H249 was subjected to the double RNase protection protocol diagrammed in Fig. 5. Thirty-microgram samples of control RNA (incubated with RNase inhibitor instead of RNase A), RNase A-treated RNA, and RNA heated to 90°C for 3 min before RNase A treatment (heated) were annealed to the sense- and antisense-specific probes used in the experiments whose results are shown in Fig. 3, corresponding to the sequences between the *Pvu*II and *Xba*I sites that bracket exon 1. To demonstrate the specificity of the probes, they were also annealed to 30 μg of yeast tRNA and RNA isolated from a c-*myc*-expressing SCLC line, H146. The probes were of equivalent sizes (sense, 805 bases; antisense, 812 bases, including the polylinker) and similar specific activities. Adjacent lanes on a single gel are shown. The numbers on the right indicate molecular sizes in bases.

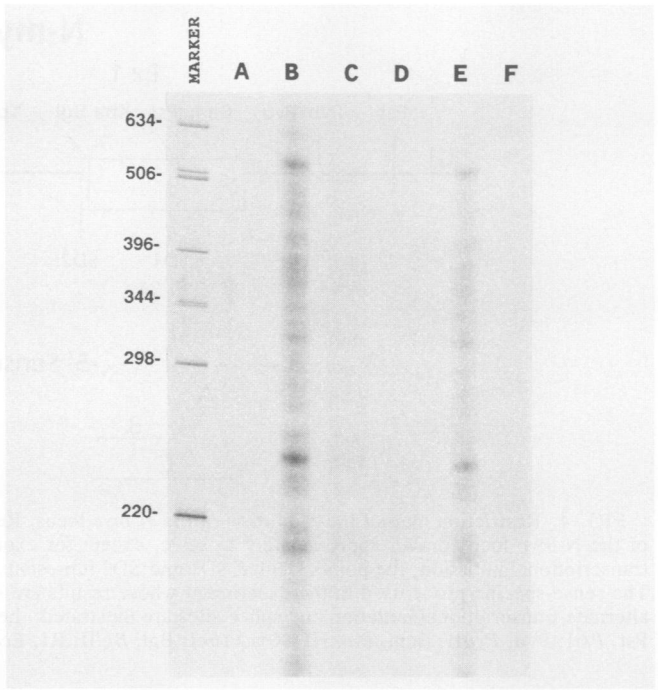


FIG. 7. RNase III digestion of RNA-RNA duplexes. Thirty-microgram samples of purified RNase A-resistant RNA from cell line H249 were digested with increasing amounts of RNase III or retreated with RNase A, annealed to the antisense-specific probe used for Fig. 3 and 6, and subjected to analytic RNase protection. Samples (30 μg) of tRNA and RNase A-resistant RNA isolated from H146, a c-*myc*-producing cell line, were used to define probe specificity. The lanes contained the following samples: A, tRNA; B, H249 RNase A-resistant RNA, untreated; C, H249 RNase A-resistant RNA treated with 0.9 μg of RNase III; D, H249 RNase A-resistant RNA treated with 1.8 μg of RNase III; E, H249 RNase A-resistant RNA retreated with 10 μg of RNase A; F, untreated H146 RNase A-resistant RNA. The numbers on the left indicate molecular sizes in bases.

sense probe from the *Eco*RI site in intron 1 to the *Bgl*II site in exon 1 and annealed the same RNA to this probe. The sense panel of Fig. 8 shows the results of the RNase protection assay. In the RNase-treated sample of RNA, there was a major band at approximately 420 bp, along with several other bands that correspond in size to the RNase A-resistant antisense species, indicating that these species exist in a duplex. In the control specimen, the band at approximately 600 bp represents the full-length probe (minus polylinker sequences), which detected a sense species which retained intron 1. This band was completely lost after RNase A digestion, indicating that the sequences between the major site of antisense initiation and the *Eco*RI site are poorly represented in duplex molecules.

**Duplex formation is, in part, determined by the initiation site of the sense RNA.** We were curious as to why, when sense N-*myc* RNA was analyzed, the complex pattern caused by multiple initiation and splice donor sites seen in the control RNA was simplified after RNase digestion (Fig. 6). Theoretically, any of the sense species generated from the multiple initiation sites should have sequences complementary to those in the primary antisense transcript and, therefore, could potentially form a duplex; the fact that this does not happen implies some selectivity in duplex formation or stability. Knowing that in procaryotic systems the precise structure at the 5′ end of the RNA molecules involved in duplex formation is critical to formation of that duplex (18), we decided to investigate whether the 5′ end of the N-*myc* RNA was important in duplex formation. To do this, we transcribed a cRNA probe from the *Pst*I site in exon

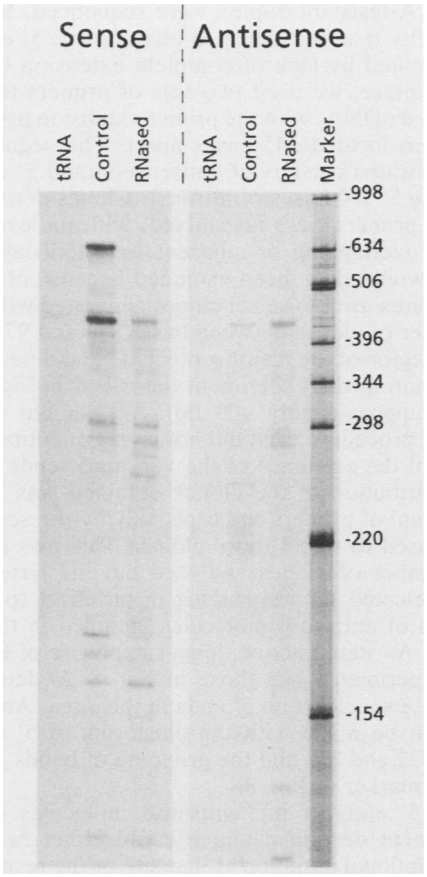


FIG. 8. Detection of the initiation sites for antisense transcription and of RNA-RNA duplexes corresponding to intron 1 sequences. A 1,097-base antisense-specific probe (antisense panel) was transcribed from the *Bgl*II site in exon 1 to the *Xho*I site in exon 2 (Fig. 4). The probe was annealed to 30-μg samples of yeast tRNA, control RNA from line H249 (incubated with RNase inhibitor), and RNase A-treated RNA (RNased), and analytic RNase protection was performed. The control lane contained the multiple transcriptional initiation sites for the antisense transcripts, with the major one at approximately 420 bases 3′ of the *Bgl*II site. The RNase-treated lane contained the corresponding RNase-resistant species. The sense panel illustrates the results of annealing the same RNA samples to a 606-base (656 bases with the polylinker) sense-specific probe transcribed from the *Eco*RI site in intron 1 to the *Bgl*II site in exon 1. The numbers on the right indicate molecular sizes in bases.

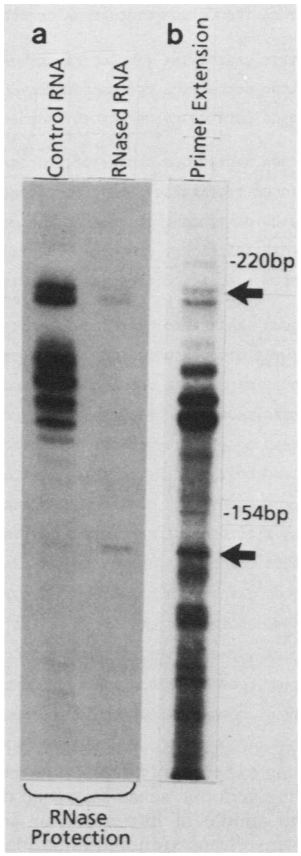


FIG. 9. Analysis of the 5′ end of the sense strand of the RNA duplex. A 474-base sense-specific probe designed to display the 5′ transcriptional initiation sites was transcribed from the *Pst*I site within exon 1 to the *Pvu*II site 5′ of exon 1 (Fig. 4). Panel a demonstrates the results of annealing this probe to control RNA and RNase-treated RNA from line H249. The initiation sites preferred in duplexes are seen in the RNase-treated (RNased) lane. Primer extension (b) confirmed these as initiation sites. The oligonucleotide primer used, whose 5′ end is the *Pst*I site, was CTGCAGGGGATC CGGAGGCGATTCTG.

1 to the *Pvu*II site 5′ of the gene (Fig. 4) to display the most 5′ major sense initiation sites in detail. Figure 9a illustrates the results obtained by annealing control RNA and RNase-resistant RNA from H249 to this probe in an RNase protection assay. Only three of the several species presumably representing independent initiation sites were present in the duplex molecules: the major most 5′ doublet and a species representing a minor initiation site approximately 140 bp from the *Pst*I site (arrows). To confirm that these species actually corresponded to initiation sites, we used an oligonucleotide primer whose 5′ border is the *Pst*I site to perform a primer extension of H249 RNA prepared by the guanidine isothiocyanate method. Our results (Fig. 9b) show that these species in fact represent independent initiation sites, demonstrating that duplex formation is, at least in part, dependent on the precise site of sense RNA initiation. These initiation sites correspond well to those previously mapped (23, 38). It is noteworthy that the distance between an initiation site included in a duplex and one that is completely RNase sensitive was as little as 17 bases (the distance between the doublet and the grouping of tightly spaced bands immediately below it).

**Cloning of RNase A-resistant RNA duplexes.** Both to confirm our analysis of the initiation sites for antisense transcription and to define precisely the extents of the RNA-RNA duplexes, we used PCR-mediated amplification of cDNA (17) to clone each strand of the sense-antisense duplex. Sequences within duplexes, as determined by the RNase protection experiments described above were used to synthesize oligonucleotide primers specific for hybridization to the sense and antisense RNAs. These primers were used to generate overlapping orientation-specific cDNA from each RNA strand within the RNase A-resistant duplex. The 3′ end of the cDNA was tailed with dATP, both to tag the 5′ end of the RNA strands and to facilitate PCR amplification and cloning. The cDNA was amplified by PCR with an oligo(dT) primer-adapter for hybridization to the dA tail, and an orientation-specific primer located 3′ to the primer which was used to generate the cDNA (to enhance stringency).

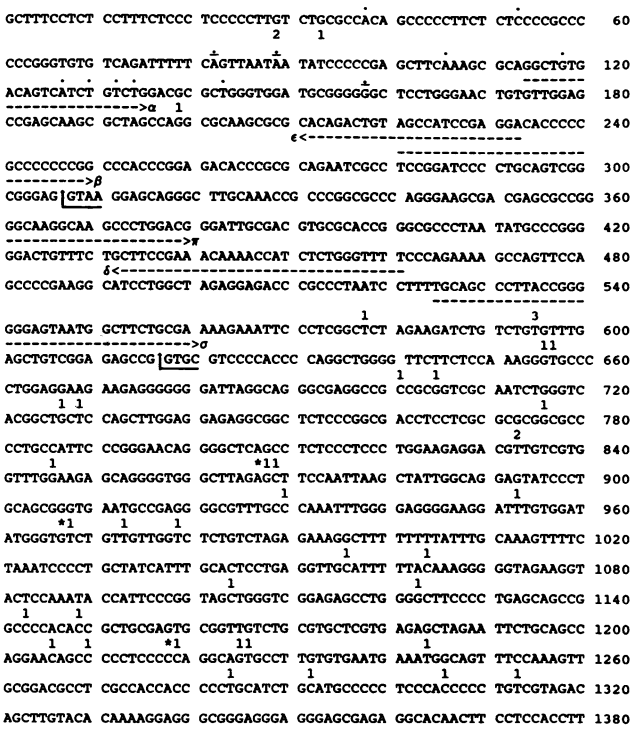

FIG. 10. Cloning of RNase A-resistant RNA duplexes. The results of the cloning experiments detailed in the text are displayed superimposed on the genomic sense sequence of N-*myc* (39), from the 5′ flank to the middle of intron 1. For additional frames of reference, the 5′ sense transcriptional initiation sites described by Stanton and Bishop (38) are indicated by dots above the appropriate bases, and the alternative splice donors determined by the same workers are indicated by the symbol ↑ within the sequence. The sense transcriptional initiation sites preferentially included in RNA duplexes (illustrated in Fig. 9) are indicated by underlined dots. The locations of the orientation-specific primers used for amplification are indicated by the arrows above the sequence (5′ − − → 3′). The nested primers used for amplification of the sense strand are primers δ and ε (the actual sequence is the complement of that illustrated). The two primer pairs used for amplification of the antisense strand are α-β, and π-σ, respectively. The 5′-most bases of the RNA represented in the cDNA clones are indicated by numbers above the bases; the numbers indicate the numbers of independent clones that map to those bases. Numbers 5′ of base 300 represent clones derived from the sense strand of the duplex; numbers 3′ of base 300 represent clones derived from the antisense strand, and therefore, the actual 5′ base would be the complement of the one illustrated.

Figure 10 illustrates the locations of the orientation-specific primers, as well as the results of the cloning and sequencing of the amplified cDNA, superimposed upon the sense N-*myc* sequence from the 5′ flank to the middle of intron 1. All clones were sequenced 250 to 300 bp in from the 5′ end of the RNA marked by the poly(dA-dT) tail of the cDNA clone. Representative clones were sequenced completely in both orientations. Because of the well-characterized sense transcriptional initiation sites detailed by other workers (23, 38) and our ability to correlate our RNase protection and primer extension results with those sites, we sequenced only four representative clones derived from the sense strand of the duplex. The 5′ ends of three of those clones (Fig. 10, bases 89 and 93) lie within 5 bp of the doublet of initiation sites shown to be preferentially included in the RNA duplexes (Fig. 9).

Thirty-four clones derived from the antisense strand of the RNase A-resistant duplex were sequenced. Because of the possibility that our ability to identify true 5′ ends may have been limited by lack of complete extension by the reverse transcriptase, we used two sets of primers to generate the amplified cDNA, with the primer extension primers (primers α and π) located 245 bases apart. The sequencing results demonstrated clusters of antisense-strand 5′ ends scattered over the 5′ 700 bases of intron 1. Clones derived from both sets of primers were intermixed, with the exception of the clones overlapping or adjacent to amplification primer σ, which would have been excluded because of the screening procedures used to select clones generated with that primer. A cluster of clones between bases 953 and 979 corresponds to the region of the major protected band detected by RNase protection in the experiment described in Fig. 8 (antisense lanes, approximately 400 bp). We should note that the cloning procedure used did not give a quantitative representation of the existence of the various 5′ ends in vivo, since the distribution of the clones obtained was biased by the placement of primers and especially by the screening procedures used to select those clones. This bias is reflected by the number of clones analyzed beyond base 1000, which were selected for sequencing in an effort to define the 5′ borders of antisense molecules included in the longest duplexes. As stated above, longer exposure of RNase protection experiments like those in Fig. 8 do demonstrate rare antisense species with 5′ ends in this area. Another correlation can be made with the clustering of 5′ ends between bases 832 and 879 and the grouping of bands just below the 298-bp marker in Fig. 8.

The 5′ ends of the antisense molecules cloned in the experiment described above could either be the result of transcriptional initiation at that site or the result of RNase A cleavage of a longer, partially duplexed molecule. We believe that many of the 5′ ends identified represent true transcriptional initiation sites because, as illustrated in Fig. 6 and 8, the qualitative distribution of antisense species did not change significantly after RNase treatment (also suggesting that most stable antisense molecules exist in a duplex). Furthermore, of the 34 clones analyzed, 27 began with a purine, most of which were adenosine residues, which would be expected from the pattern of eucaryotic transcriptional initiation. However, to address this question directly, we derived antisense-specific cDNA clones from total cellular RNA (isolated with guanidine isothiocyanate) by the procedure detailed above. We sequenced three representative clones, whose 5′ ends are denoted by numbers preceded by asterisks in Fig. 10. The clone that mapped to nucleotide 968 is consistent with the major transcriptional initiation site identified in Fig. 8 (antisense control lane, approximately 400 bp); the clone that mapped to nucleotide 868 corresponds to the minor initiation sites seen just below the 298-bp marker. The clone that mapped to nucleotide 1219 corresponds to a minor initiation site seen on long exposures of RNase protection experiments, such as that shown in Fig. 8. Notably, these 5′ ends fall into clusters identified by analyzing clones derived from RNase A-resistant RNA. Finally, we should comment on the fidelity of the sequences obtained. The only deviation from previously published sequences were occasional base changes, without a particular pattern and inconsistent between overlapping clones, suggesting polymerase errors during the amplification process.

## DISCUSSION

**Existence of stable antisense RNA derived from the N-*myc* locus.** We have demonstrated that antisense transcription

results in the formation of two populations of stable antisense transcripts derived from the N-*myc* locus. The polyadenylated population is extremely rare (<1% of the amount of polyadenylated sense RNA) and on Northern blotting migrates as 1.0- and 1.8-kb species which appear to be composed of multiple exons whose coding capacity is not known. The nonpolyadenylated antisense RNA is more abundant (approximately 5% of the total sense RNA) but heterogeneous in size, with a broad distribution just under 1 kb on Northern blotting. Transcription of these antisense RNAs initiates at multiple sites predominately within intron 1, as shown by RNase protection and cDNA cloning, which correlates well with the results of our nuclear runoff in vitro transcription assays. The major site of transcriptional initiation lies in an A+T-rich region near the *Xba*I site (coordinate 985 in Fig. 10) in the middle of intron 1. There is no TATA box associated with this or any of the other minor 5′ or 3′ initiation sites. The only recognizable potential regulatory sequences related to the major initiation site are three putative octamer (36) motifs (7 of 8 bp matches, each with the hexanucleotide core TTTGCA preserved) within 60 to 70 bases of the start sites.

The pattern of antisense transcriptional initiation, with multiple start sites scattered over a few hundred bases, is strikingly similar to that of sense initiation. The N-*myc* gene falls into the category of polymerase II-transcribed genes which lack a TATA box, have multiple potential binding sites for transcription factor SP1 (14), and initiate transcription at multiple sites scattered over relatively large regions (see reference 37 for a discussion of eucaryotic transcriptional initiation). In fact, not only are there four putative SP1-binding sites in the region of the sense transcription initiation sites, but there is also one within intron 1 and there are five within exon 2. It therefore seems plausible that the entire region functions as a bidirectional promoter of transcription in the sense and antisense orientations. A fact in favor of this hypothesis is that transcription in both orientations appears to be coordinately regulated. For example, despite widely varying rates of transcription among different cell lines, the ratio of sense to antisense transcription across exon 1 remains constant (Fig. 1). Furthermore, upon differentiation of neuroblastoma cells with retinoic acid, which decreases N-*myc* expression (40), the levels of stable sense and antisense RNAs fall in parallel (G. W. Krystal, unpublished data).

There is some precedence in higher eucaryotes for bidirectional promoters that give rise to opposite-strand transcripts. For instance, the mouse α1(IV) and α2(IV) collagen genes are arranged in a head-to-head fashion separated by 130 nucleotides and regulated by a common bidirectional promoter and shared enhancer (7). Another example of a bidirectional promoter occurs in the mammalian *dhfr* locus, which, like N-*myc* and the type IV collagen genes, lacks a TATA box and has multiple consensus SP1-binding sites in a highly G+C-rich region. The locus gives rise to small poly(A)$^-$ (15, 27) and multiexon poly(A)$^+$ opposite-strand transcripts (27, 30), some of which intriguingly contain sequences complementary to sense *dhfr* transcripts derived from a minor upstream promoter.

**RNA-RNA duplexes exist in vivo between specific N-*myc* transcripts and antisense transcripts derived from the N-*myc* locus.** By using a double RNase protection technique, we have demonstrated the existence of RNase A-resistant transcripts corresponding to sense and antisense sequences derived from the 5′ portion of the N-*myc* gene. We believe that the RNase resistance does not arise from self-protection by the individual transcripts, because if that were the case, one would expect the large excess of sense transcripts to be reflected in the amount of RNase-resistant material; instead, we invariably detected equivalent amounts of resistant sense and antisense RNAs. Furthermore, when excess radioactively labeled probes were added to lysates, no protection was seen, effectively ruling out both self-protection by these sequences and in vitro duplex formation as explanations of our results. We do not believe the protection results from the formation of RNA-DNA hybrids, since we routinely treat our cytoplasmic supernatants with DNase and treatment with RNase H before analytic RNase protection does not alter the results (data not shown). Although protein factors may play a role in formation (or preventing formation) of the RNase-resistant duplexes, they cannot account for the RNase resistance because purified RNA (after proteinase K digestion and phenol-chloroform extraction) remained RNase A resistant. Furthermore, the purified RNase A-resistant RNA was sensitive to a double-strand-specific RNase, RNase III. The correspondence in length of the RNase A-resistant sense and antisense RNAs, along with the RNase III sensitivity, strongly suggests that the only explanation for the RNase A resistance is formation of an RNA-RNA duplex between complementary sense and antisense transcripts. Brief heat denaturation before RNase digestion, which resulted in complete loss of signal with either sense- or antisense-specific probes, supports this conclusion. Finally, by using a strategy designed to clone each strand of the RNA duplex from RNase A-resistant RNA independently in an orientation-specific manner, we demonstrated overlapping sense- and antisense-specific clones. We believe that the evidence described above conclusively demonstrates the existence of an RNA-RNA duplex between sense and antisense transcripts derived from the N-*myc* locus. Mixing experiments demonstrated that the RNA-RNA duplexes were not formed after the cells were lysed.

There is some precedence for the existence of RNA-RNA duplexes in eucaryotic cells. Approximately 2 to 5% of the heterogeneous nuclear RNA isolated by traditional methods is RNase resistant (8); these RNA-RNA duplexes can be visualized by electron microscopy (16). $R_0t$ analysis suggested that most of the duplex molecules were intramolecular duplexes or intermolecular duplexes between reiterated sequences, although the gross nature of the analysis did not exclude the possibility that single-copy sequences made some contribution to the total RNase-resistant material. Besides the demonstration of double-stranded heterogeneous nuclear RNA, other evidence which addresses the presence of RNA duplexes is the number of enzymatic activities which have been described that unwind exogenously supplied double-stranded RNA. In addition to the activities associated with splicing (24, 35) and translation (26, 33), a nuclear activity exists in *Xenopus* embryos and several mammalian cell lines which not only unwinds long, perfect RNA duplexes (2, 34, 42) but inhibits their reformation through the conversion of adenosine to inosine residues (3). Recently, Kimelman and Kirschner have demonstrated that an antisense transcript containing complementary sequences to exon 3 of basic fibroblast growth factor (bFGF) exists in *Xenopus* oocytes (21). Those workers also demonstrated that following germinal vesicle breakdown, when the cytoplasmic RNA would be exposed to nuclear double-stranded RNA unwindase, bFGF mRNA undergoes extensive conversion of its adenosine residues to inosine in exon 3, which results in A-to-G conversions after cDNA cloning. This appears to be indirect evidence for the existence of an

RNA-RNA duplex between bFGF and its overlapping antisense RNA. Interestingly, this covalent modification of bFGF RNA appears to be temporally related to rapid degradation of bFGF mRNA, suggesting a role for the RNA duplex in the regulation of mRNA stability. The RNA-RNA duplexes that cover the 5′ portion of the N-*myc* gene seem to be a potential substrate for the double-stranded RNA unwindase enzyme. We did not detect any A-to-G conversions in sequencing our cDNA clones, but we would not expect to find any, since they were derived from RNase-resistant RNA, and following activity of the RNA unwindase enzyme, double-stranded RNA becomes sensitive to RNase A (3).

**Duplex formation is, in part, determined by the transcriptional initiation site.** After RNase treatment, we noted differential resistance among the multiple forms of N-*myc* sense RNA, suggesting that not all forms engaged in duplex formation with complementary antisense RNA equally (Fig. 6). Mapping of the 5′ ends of the sense RNA within duplexes and comparing them with transcriptional initiation sites authenticated by primer extension (Fig. 9) resulted in the conclusion that duplex formation is, in part, determined by the structure of the 5′ end, with as little as 17 bases separating the 5′ ends of RNase-resistant and RNase-sensitive RNAs. These findings were confirmed by cloning the 5′ ends of the sense strand of the RNA duplex (Fig. 10). The reason for the preference for specific 5′ ends within duplexes is not clear.

**Potential functions of N-*myc* RNA-RNA duplexes.** Our data suggest that the level of translatable N-*myc* mRNA is not regulated by the amount of antisense RNA available to form a duplex because the transcriptional rate and the steady-state levels of sense and antisense transcripts have always been proportionately linked under all of the physiologic circumstances we have explored. Our data do, however, suggest a potential function for the RNA duplexes simply on a structural basis. Virtually all of the duplexes cloned extend across both exon 1 splice donor sites, potentially protecting them from the splicing enzyme complex (Fig. 10). Konarska et al., on the basis of in vitro data, have already speculated that intermolecular RNA duplexes may regulate splicing in such a fashion (24). Duplex formation, therefore, may play a role in regulating the splicing patterns of the specific N-*myc* RNAs capable of forming a duplex. We have identified a 4.2-kb polyadenylated cytoplasmic RNA corresponding to an N-*myc* RNA that retains intron 1 with intron 2 spliced out. The significance of this mRNA is not clear, but it is worth noting that a large proportion of L-*myc* mRNAs retain intron 1 (19), which is required to generate an alternative form of the L-*myc* protein (12).

We can speculate on other potential functions of the RNA duplexes. Duplex formation may play a role in regulating mRNA stability, as demonstrated for *Xenopus* bFGF. The duplexes, which extend across the exon 1-intron 1 interface, may also play a role in abrogating a block to transcriptional elongation, which is present in this region during c-*myc* and L-*myc* transcription, but not during N-*myc* transcription (Fig. 1; [25]). Finally, if the N-*myc* duplexes are a substrate for double-stranded nuclear RNA unwindase, it is possible that the coding potential of the specific RNAs involved could be altered through the adenosine-to-inosine conversions.

In summary, we have demonstrated the existence of stable antisense transcripts from the N-*myc* locus which exist in an RNA-RNA duplex with specific sense transcripts. Duplexes between sense and antisense bFGF transcripts have also been indirectly demonstrated, and so it is possible that these structures play an important role in regulating the processing and stability of eucaryotic mRNA. By using the techniques we have described to isolate and characterize these RNA duplexes, it is possible to design experiments to determine their functional significance in the regulation of these and other genes.

## ACKNOWLEDGMENTS

We thank John Minna for many helpful discussions and for contributing the genomic phage clone containing the N-*myc* locus. We also thank Donald Court, Frederick Cancer Research Facility, National Cancer Institute, for his generous gift of purified RNase III and James Way for excellent technical assistance during the early part of this work.

This work was supported in part by grants from the American Cancer Society (αIN 105-N) and the Department of Veterans Affairs to G.K.

## LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligo-thymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1413.
2. **Bass, B. L., and H. Weintraub.** 1987. A developmentally regulated activity that unwinds RNA duplexes. Cell **48**:607–613.
3. **Bass, B. L., and H. Weintraub.** 1988. An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell **55**:1089–1098.
4. **Battey, J., C. Moulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder.** 1983. The human c-*myc* oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell **34**:779–787.
5. **Bentley, D. L., and M. Groudine.** 1986. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL-60 cells. Nature (London) **321**:702–706.
6. **Bentley, D. L., and M. Groudine.** 1986. Novel promoter upstream of the human c-*myc* gene and regulation of c-*myc* expression in B-cell lymphomas. Mol. Cell. Biol. **6**:3481–3489.
7. **Burbelo, P. B., G. R. Martin, and Y. Yamada.** 1988. α1(IV) and α2(IV) collagen genes are regulated by a bidirectional promoter and a shared enhancer. Proc. Natl. Acad. Sci. USA **85**:9679–9682.
8. **Calvert, J. P., and T. Pederson.** 1977. Secondary structure of hnRNA: two classes of double-stranded RNA in native ribonucleoprotein. Proc. Natl. Acad. Sci. USA **74**:3705–3709.
9. **Chirgwin, J. A., A. Aeybyle, R. McDonald, and W. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources rich in ribonuclease. Biochemistry **18**:5294–5299.
10. **Davis, L. G., M. D. Dibner, and J. F. Battey.** 1986. Basic methods in molecular biology. Elsevier Science Publishing, Inc., New York.
11. **Dean, M., R. B. Kent, and G. E. Sonenshein.** 1983. Transcriptional activation of immunoglobulin α heavy chain genes by translocation of the c-*myc* oncogene. Nature (London) **305**:443–446.
12. **De Greve, J., J. Battey, J. Fedorko, M. Birrer, G. Evan, F. Kaye, E. Sausville, and J. Minna.** 1988. The human L-*myc* gene encodes multiple nuclear phosphoproteins from alternatively processed mRNAs. Mol. Cell. Biol. **8**:4381–4388.
13. **DePinho, R. A., K. S. Hatton, A. Tesfaye, G. D. Yancopoulos, and F. W. Alt.** 1987. The human *myc* gene family: structure and activity of L-*myc* and an L-*myc* pseudogene. Genes Dev. **1**:1311–1326.
14. **Dynan, W. S., and R. Tijan.** 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35**:79–87.
15. **Farnham, P. J., J. M. Abrams, and R. T. Schimke.** 1985. Opposite-strand RNAs from the 5′ flanking region of the mouse dihydrofolate reductase gene. Proc. Natl. Acad. Sci. USA **82**:3978–3982.
16. **Federoff, N., P. K. Wellauer, and R. Wall.** 1977. Intermolecular duplexes in heterogeneous nuclear RNA from HeLa cells. Cell **10**:597–610.
17. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid

production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.

18. **Green, P. J., O. Pines, and M. Inouye.** 1986. The role of antisense RNA in gene regulation. Annu. Rev. Biochem. **55**: 569–597.
19. **Kaye, F., J. Battey, M. Nau, B. Brooks, E. Seifter, J. DeGreve, M. Birrer, E. Sausville, and J. Minna.** 1988. Structure and expression of the human L-*myc* gene reveal a complex pattern of alternative mRNA processing. Mol. Cell. Biol. **8**:186–195.
20. **Keath, E. J., A. Kelekar, and M. D. Cole.** 1984. Transcriptional activation of the c-*myc* oncogene in mouse plasmacytomas: similar RNA levels in tumor and proliferating normal cells. Cell **37**:521–528.
21. **Kimelman, D., and M. W. Kirschner.** 1989. An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in Xenopus oocytes. Cell **59**:687–696.
22. **Kindy, M. S., J. E. McCormack, A. J. Buckler, R. A. Levine, and G. E. Sonenshein.** 1987. Independent regulation of transcription of the two strands of the c-*myc* gene. Mol. Cell. Biol. **7**:2857–2862.
23. **Kohl, N. E., E. Legouy, R. A. DePinho, P. D. Nisen, R. K. Smith, C. E. Gee, and F. W. Alt.** 1986. Human N-*myc* is closely related in organization and nucleotide sequence to c-*myc*. Nature (London) **319**:73–77.
24. **Konarska, M. M., R. A. Padgett, and P. A. Sharp.** 1985. Trans splicing of mRNA precursors *in vitro*. Cell **42**:165–171.
25. **Krystal, G., M. Birrer, J. Way, M. Nau, E. Sausville, C. Thompson, J. Minna, and J. Battey.** 1988. Multiple mechanisms for transcriptional regulation of the *myc* gene family in small-cell lung cancer. Mol. Cell. Biol. **8**:3373–3381.
26. **Liebhaber, S. A., F. E. Cash, and S. H. Shakin.** 1984. Translationally associated helix-destabilizing activity in rabbit reticulocyte lysate. J. Biol. Chem. **259**:15597–15602.
27. **Linton, J. P., J.-Y. J. Yen, E. Selby, Z. Chen, J. M. Chinsky, K. Liu, R. E. Kellems, and G. F. Crouse.** 1989. Dual bidirectional promoters at the mouse *dhfr* locus: cloning and characterization of two mRNA classes of the divergently transcribed *Rep-1* gene. Mol. Cell. Biol. **9**:3058–3072.
28. **Maniatis, T., and R. Reed.** 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. Nature (London) **325**:673–678.
29. **McCarthy, T. L., E. Siegel, B. Mroczkowski, and S. M. Heywood.** 1983. Characterization of translational control ribonucleic acid isolated from embryonic chicken muscle. Biochemistry **22**:935–941.
30. **Mitchell, P. J., A. M. Carothers, J. H. Han, J. D. Harding, E. Kas, L. Venolia, and L. A. Chasin.** 1986. Multiple transcription start sites, DNase I-hypersensitive sites, and an opposite-strand exon in the 5′ region of the CHO *dhfr* gene. Mol. Cell. Biol. **6**:425–440.
31. **Nau, M. M., B. J. Brooks, D. N. Carney, A. F. Gazdar, J. F. Battey, E. Sausville, and J. D. Minna.** 1986. Human small cell lung cancers show amplification and expression of the N-*myc* gene. Proc. Natl. Acad. Sci. USA **83**:1092–1096.
32. **Nepveu, A., and K. B. Marcu.** 1986. Intragenic pausing and anti-sense transcription within the murine c-*myc* locus. EMBO J. **5**:2859–2965.
33. **Ray, B. K., T. C. Lawson, J. C. Kramer, M. H. Cladaras, J. A. Grifo, R. D. Abramson, W. C. Merrick, and R. E. Thatch.** 1985. ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. J. Biol. Chem. **260**:7651–7658.
34. **Rebagliati, M. R., and D. A. Melton.** 1987. Antisense RNA injections in fertilized frog eggs reveals an RNA duplex unwinding activity. Cell **48**:599–605.
35. **Ruskin, B., and M. R. Green.** 1985. Specific and stable intron-factor interactions are established early during *in vitro* pre-mRNA splicing. Cell **43**:131–142.
36. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
37. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57**:103–113.
38. **Stanton, L. W., and J. M. Bishop.** 1987. Alternative processing of RNA transcribed from N-*myc*. Mol. Cell. Biol. **7**:4266–4272.
39. **Stanton, L. W., M. Schwab, and J. M. Bishop.** 1986. Nucleotide sequence of the human N-*myc* gene. Proc. Natl. Acad. Sci. USA **83**:1772–1776.
40. **Thiele, C. J., C. P. Reynolds, and M. A. Israel.** 1985. Decreased expression of N-*myc* precedes retinoic acid-induced morphological differentiation of human neuroblastoma. Nature (London) **313**:404–406.
41. **van der Krol, A. R., J. N. M. Mol, and A. R. Stuitje.** 1988. Modulation of eukaryotic gene expression by complementary RNA or DNA sequences. BioTechniques **6**:958–976.
42. **Wagner, R. W., and K. Nishikura.** 1988. Cell cycle expression of RNA duplex unwindase activity in mammalian cells. Mol. Cell. Biol. **8**:770–777.
43. **Winter, E., F. Yamamoto, C. Almoguera, and M. Perucho.** 1985. A method to detect and characterize point mutations in transcribed genes: amplification and overexpression of the mutant c-Ki-*ras* allele in human tumor cells. Proc. Natl. Acad. Sci. USA **82**:7575–7579.