Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3204–3208, May 1986
Biochemistry

Murine *myb* protooncogene mRNA: cDNA sequence and evidence for 5′ heterogeneity

(mRNA structure/B-lymphocytes)

TIMOTHY P. BENDER AND W. MICHAEL KUEHL

NCI–Navy Medical Oncology Branch, National Cancer Institute, National Institutes of Health and Naval Hospital, Bethesda, MD 20814-2015

Communicated by Oscar L. Miller, Jr., January 3, 1986

ABSTRACT **We have sequenced two overlapping cDNA clones from a murine pre-B cell library to generate a composite sequence that includes 3413 bases of the murine c-myb mRNA. There is a single long open reading frame, beginning at the first base of this sequence, and continuing from the first methionine codon at nucleotide 265 to a TGA termination codon at nucleotide 2173. The predicted murine translation product contains 636 amino acid residues and is about 71 kDa long, which is in good agreement with the 75-kDa molecular size determined for the avian c-myb protein. The murine c-myb protein shows a striking 82% amino acid homology in the region (amino acids 71–444) where it can be compared to the published avian c-*myb* gene sequence. S1 nuclease protection analysis indicates extreme heterogeneity at the 5′ end of steady-state murine c-myb mRNA.**

The *myb* protooncogene, which appears to encode a nuclear DNA binding protein (1–3), is expressed predominantly in hematopoietic cells (4–9). Expression of c-*myb* occurs in both normal and transformed cells derived from all species and hematopoietic lineages examined. Moreover, in each hematopoietic lineage (erythroid, myeloid, and lymphoid) examined, c-myb mRNA expression is much greater in immature cells than in more differentiated cells (7–11). The specificity of c-*myb* expression in hematopoietic cells correlates with the specific transformation of hematopoietic cells by replication-defective avian acute leukemia viruses that express an internal portion of the avian c-*myb* gene as a fusion product (12–18). Avian myeloblastosis virus (AMV) causes myeloblastic or monocytic leukemia in chickens. Avian erythroblastosis virus E26, which contains a smaller internal segment of the avian c-*myb* gene than AMV but also contains a portion of the avian c-*ets* gene, fused to the 3′ end of c-*myb*, causes erythroblastic as well as myeloblastic or monocytic leukemia in chickens. Changes in endogenous c-*myb* genes have also been observed in both human and murine hematopoietic tumors. First, at least four independent murine plasmacytoid lymphosarcomas are associated with insertion of a defective Moloney murine leukemia virus into a c-*myb* gene and with expression of an abnormal c-*myb* transcription product (19). Second, amplification of c-*myb* has been reported in a single case of human myeloid leukemia (20).

We began to determine the structure and transcription unit of the murine c-myb mRNA as a first step in attempting to understand the function of c-*myb* in normal and transformed hematopoietic cells. In this report, we present the nucleotide sequence of murine c-myb mRNA. We also provide evidence suggesting that there is extreme heterogeneity at the 5′ end of murine c-myb mRNA.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

cDNA Cloning and Isolation of c-myb cDNA Clones. Poly(A)-positive total cellular RNA was isolated from the 70Z/3B cell line (21–23). Double-stranded cDNA was made using RNase H and DNA polymerase I for second-strand synthesis (24, 25). The cDNA was made blunt ended with T4 DNA polymerase, methylated, and ligated to *Eco*RI linkers (Collaborative Research, Lexington, MA). After digestion with *Eco*RI the cDNA was fractionated on a 5% polyacrylamide gel and cDNA >500 base pairs (bp) was ligated into lambda gt10 and gt11 phage cloning vectors as described by Young and Davis (26). Phage plaques transferred to nitrocellulose filters were screened using the *Kpn* I–*Xba* I fragment from AMV (ref. 15; provided by E. P. Reddy). The filters were washed at 45°C in 0.1× SSC/0.1% $NaDodSO_4$ (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). Approximately 250,000 recombinant plaques were screened from each library.

Nucleotide Sequence Analysis. The DNA sequence was determined by isolation and subcloning of specific restriction fragments into M13mp10 and mp11 vectors essentially as described by Bethesda Research Laboratory. The entire protein coding region and >95% of the composite message described were sequenced from both strands, and all sites used for cloning were crossed in the sequencing.

S1 Nuclease Analysis. Total cellular RNA was used for protection studies. Single-stranded DNA probes complementary to transcribed sequences were prepared by primer extension of M13 clones as described (27) in the presence of [^{32}P]dATP. After annealing and S1 digestion, protected species were fractionated on 8 M urea/5% polyacrylamide gels.

RESULTS

RNA blot hybridization analysis of mRNA from various murine pre-B cell lymphomas shows at least two forms of c-myb mRNA, a major species of approximately 3.8 kilobases (kb) and a minor species of approximately 4.2 kb (ref. 28 and T.P.B., unpublished data). Using mRNA from the 70Z/3B murine pre-B cell lymphoma line (21), which expresses high levels of these two species of c-myb mRNA, two cDNA libraries (λgt10 and λgt11) were prepared. Five clones from the gt11 library and four clones from the gt10 library were isolated by screening with an avian v-myb probe. One clone, gt10-10, contains approximately 3.2 kb of cDNA insert. Restriction mapping of the other eight clones indicates that, although all of the clones overlap the gt10-10 clone, a second clone, gt11-15, contains about 300 bp that extend to the 5′ side of the gt10-10 clone (Fig. 1).

Structure of c-myb mRNA. The gt10-10 and gt11-15 cDNA clones were sequenced. The composite c-myb mRNA se-

Abbreviations: kb, kilobase(s); bp, base pair(s); AMV, avian myeloblastosis virus.

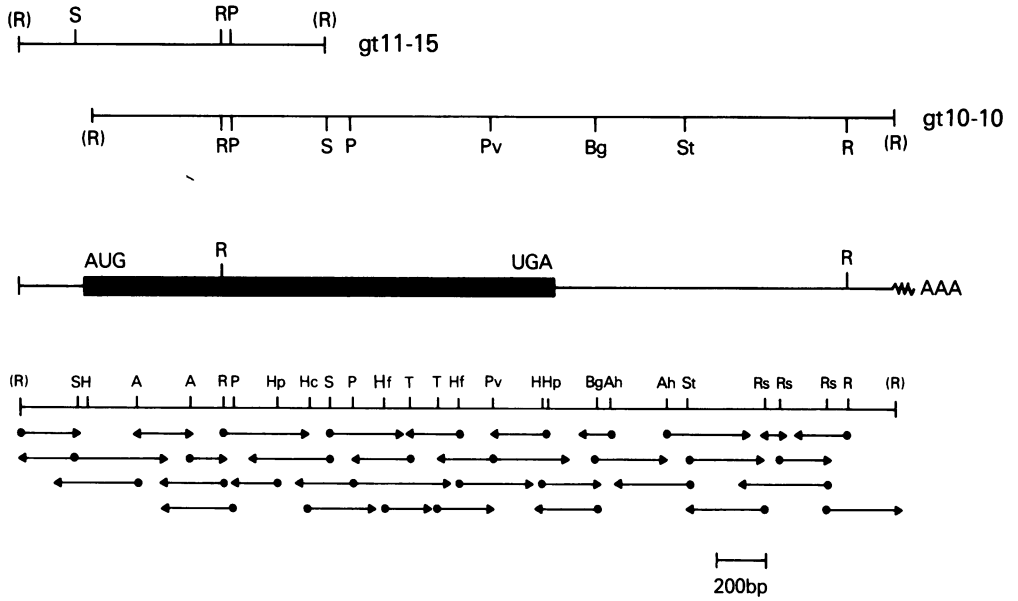

FIG. 1. Restriction map and sequencing strategy for the c-myb cDNA clones gt10-10 and gt11-15. A diagrammatic figure of the proposed c-myb mRNA is provided with relevant features for comparison with the cDNA clones. Shown in bold line is the putative c-myb protein coding region, translation initiation and termination codons, internal *Eco*RI restriction sites, and poly(A) tail. Abbreviations for restriction cleavage sites are as follows: R, *Eco*RI; S, *Sma* I; H, *Hae* III; P, *Pst* I; Pv, *Pvu* II; Bg, *Bgl* II; St, *Sst* I; A, *Alu* I; Hc, *Hin*cII; Hp, *Hpa* II; Hf, *Hin*fII; T, *Taq* I; Ah, *Aha* III; Rs, *Rsa* I. Not all restriction sites are shown in the sequencing strategy.

quence includes 3413 nucleotides, not including 19 adenine residues that are presumably derived from the poly(A) tail (Fig. 2). There is a single long open reading frame that begins at the first base of this sequence and extends to a TGA termination codon at nucleotide 2173. The first methionine codon in the long open reading frame is found at position 265. Since this methionine codon occurs in a stretch of sequence that fits the consensus eukaryotic translation start sequence, it very likely represents the translation initiation site (30). The methionine codon at position 26 would not serve as a translation initiation site because it does not occur in a context that fits the consensus eukaryotic translation start sequence and is also followed by a TAA termination codon in frame at position 53 (Fig. 2). If the methionine at position 265 is the correct site of initiation of translation, the primary translation product would contain 636 amino acid residues and have a size of about 71 kDa, in good agreement with the 75-kDa molecular size determined for the avian c-myb protein (18). Twenty-six residues upstream from the 3′ stretch of adenine residues, at residue 3388, is an AATAAA consensus polyadenylylation signal (31). When additional cDNA clones were selected by screening with a 3′ *Sst* I/*Eco*RI (nucleotides 2678–3349) probe, three of the four new clones identified terminated just after the polyadenylylation signal at position 3388, and one clone terminated substantially before this signal. A second AATAAA sequence is noted at position 2468, but at present we have no evidence that this site is used as a polyadenylylation signal. The c-myb mRNA includes 1241 nucleotides of 3′-untranslated sequence.

Detection of Genomic c-*myb* Sequences. BALB/c genomic DNA was digested with either *Eco*RI or *Bam*HI endonucleases, fractionated on 0.8% agarose gels, and transferred to nitrocellulose filters. As shown in Fig. 3, murine cDNA probes identified the following eight *Eco*RI fragments: 7.8- and 4.2-kb fragments with the 5′ probe I (nucleotides 1–820); 13.6-, 3.3-, 2.1-, 1.7-, and 1.4-kb fragments with probe II (nucleotides 821–3350); and a 4.0-kb fragment with the 3′ probe III (nucleotides 3351–3432). Four of these fragments (4.2, 2.1, 1.7, and 1.4 kb) are also seen when hybridized to a v-myb probe (ref. 19 and T.P.B., unpublished data).

To better define the 5′ end of the murine c-myb transcription unit, genomic c-myb clones were isolated from an embryonic BALB/c library (provided by P. Leder). A 5′ c-myb cDNA probe (nucleotides 1–232 in Fig. 2) hybridized to the 7.8-kb *Eco*RI genomic fragment but not to the 4.2-kb *Eco*RI genomic fragment (data not shown). A 1.1-kb *Bam*HI subfragment which hybridized to this 5′ c-myb cDNA probe, was isolated from the 7.8-kb genomic fragment. The 1136-bp sequence of this *Bam*HI genomic subfragment, shown in Fig. 4, includes 799 bp to the 5′ side of and 50 bp to the 3′ side of a sequence identical to nucleotides 1–287 of the murine cDNA (Fig. 2).

S1 Nuclease Protection Studies Indicate Heterogeneity at the 5′ End of c-myb mRNA. We have used the 1136-bp *Bam*HI genomic fragment for S1 nuclease protection studies. The fragment was annealed to 70Z/3B total cellular RNA, digested with S1 nuclease, and the protected products fractionated on a denaturing polyacrylamide gel. As shown in Fig. 5*A*, at least 13 protected products were detected, with approximate sizes as follows: 184, 195, 211, 224, 235, 250, 278, 306, 346, 368, 495, 680, and 970 nucleotides (as depicted in Fig. 4). A similar pattern of protected products was found with other preparations of total, total poly(A)-positive, nuclear and cytoplasmic 70Z/3B RNA as well as with 1881 and LS8.T2 pre-B-cell total cellular RNA (data not shown). We have also used a 592-base *Sma* I/*Eco*RI fragment corresponding to bases 233–824 of our composite cDNA sequence for S1 nuclease protection studies. This probe, probe B in Fig. 5*B*, contains 36 nucleotides of 5′-untranslated sequence, and 556 nucleotides of protein coding sequence. As shown in Fig. 5*B*, a single full-length protected species was detected after S1 nuclease digestion showing that all of the detected heterogeneity occurs 5′ of the translation initiation codon.

The longest product (950 nucleotides) protected by probe A (Fig. 5*A*), together with the appropriate region of the composite cDNA sequence, predicts a mRNA species of about 4.2 kb. Thus, this region may include the 5′ most genomic sequences transcribed into murine c-myb mRNA. With the smaller protected species, which range in size from about 188–680 bases, it can be predicted that they would derive from mRNA species of 3.5–4.0 kb. The major 3.8-kb mRNA species seen on RNA blots may represent an average of multiple species.

DISCUSSION

Addition of 150 adenine residues to the 3413 nucleotide murine c-myb cDNA (Fig. 2) predicts a mRNA containing 3563 nucleotides. This is somewhat shorter than the major 3.8-kb mRNA species and substantially shorter than the minor 4.2-kb mRNA species detected on RNA blots of c-myb

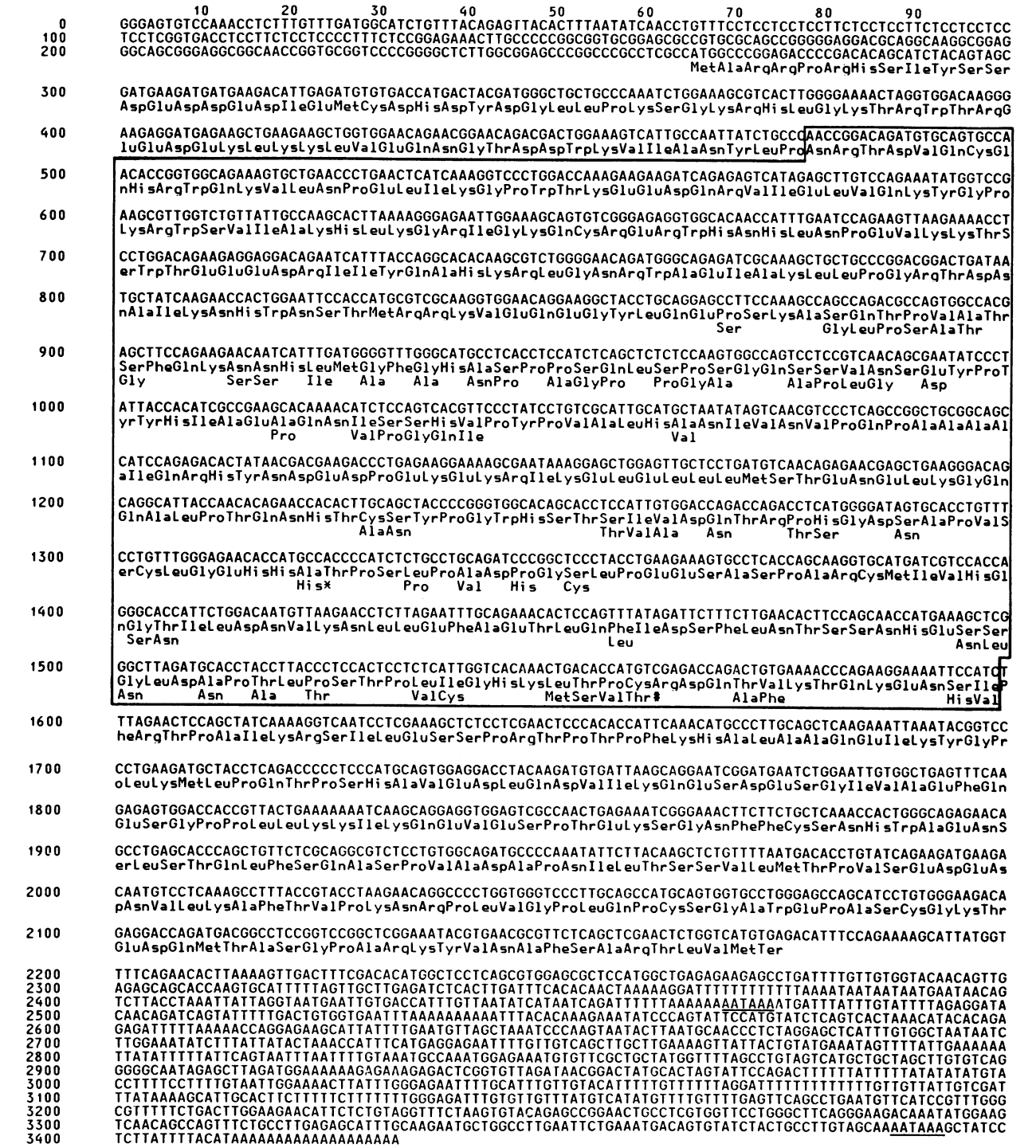

FIG. 2. Nucleotide sequence of the murine c-myb mRNA. The amino acid sequence of the single long open reading frame, beginning at the first in frame initiation codon, is shown directly below the nucleotide sequence. The boxed region indicates the available amino acid sequence of the avian c-*myb* (16). Amino acid differences between the murine and avian sequence are indicated beneath the murine amino acid sequence. Deletions in the murine sequence relative to the avian sequence are marked * (one amino acid) and # (four amino acids). The two potential AATAAA polyadenylylation signals are underlined. Discordances with the published murine c-myb mRNA sequence (ref. 29; corresponding to positions 230–3020 in our sequence) are indicated as the position in our sequence. The nucleotide difference in the published sequence and the amino acid change in the published sequence in parentheses are as follows: 251/252, insert C; 492, substitute A; 577, substitute A (Lys); 865, substitute A (Lys); 1064, substitute T (Val); 1496, substitute A (Asn); 1763, substitute G (Arg); 1788, substitute T; 1838, substitute C (Ala); 1914, substitute A; 2376, delete T; 2377, substitute A; 2497, substitute A; 2721, delete C; 2724-5, delete AA; 2726, substitute G; 2732, delete C; 2742, delete A; 2766, delete A; 2774, delete C; 2783, delete A; 2875, substitute A; 2893, substitute A; 2989/2990, insert T; 3014/3015, insert C; 3020, substitute A.

mRNA from various murine hematopoietic cells (ref. 28 and T.P.B., unpublished data). The apparent 5′ heterogeneity indicated by S1 nuclease protection studies (Figs. 4 and 5) may account for the discrepancy between the size of c-myb mRNA as determined by RNA blots and the size predicted from our composite sequence of two cDNA clones. At present we have no evidence as to whether the apparent 5′ heterogeneity is due to multiple sites of transcription initiation, to posttranscriptional cleavage at the 5′ end of the mRNA, or to alternative splicing of a 5′-leader sequence.

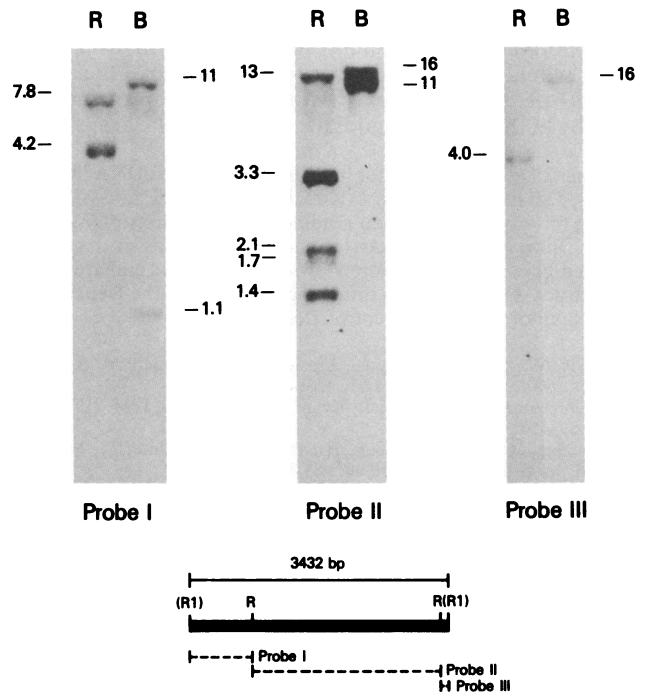

FIG. 3. Southern blot analysis of BALB/c genomic DNA. The DNA was digested with either *Eco*RI (R) or *Bam*HI (B), fractionated on 0.8% agarose gels, and transferred to nitrocellulose (32). The fragment sizes are listed to the side of each digest in kilobases. *Eco*RI sites which are the result of *Eco*RI linkers are indicated as (R1). Nick-translated probes were prepared from specific fragments as indicated above and include the following bases relative to the c-myb mRNA sequence in Fig. 2: probe I, nucleotides 1–820; probe II, nucleotides 821–3350; probe III, nucleotides 3351–3432. Filters were hybridized at 42°C and washed in 0.1× SSC/0.1% $NaDodSO_4$ at 58°C (probes I and II) or 50°C (probe III).

The published murine c-myb mRNA sequence of Gonda *et al.* (29) (corresponding to positions 230–3020 in our sequence) predicts essentially the same coding sequence but does not define either the 5′ or 3′ ends of the mRNA. As indicated in Fig. 2, it differs at 9 coding positions (resulting in six amino acid differences) and 18 noncoding positions. These differences may reflect sequencing errors, cloning artifacts, somatic mutations of c-*myb* in the cell lines, or genetic polymorphisms. We can provide some additional evidence regarding our sequence. First, all putative coding sequences and most noncoding sequences were sequenced in two directions (Fig. 1). Second, positions 472, 577, and 865 were identical for independent cDNA clones. However, two independent cDNA clones differed at position 1064, with one clone identical to the Gonda *et al.* (29) sequence. Third, our sequence lacks a *Hin*dIII site (position 577) and contains a *Pvu* II site (position 1914), which are discrepant with the sequence of Gonda *et al.* (29). Both discrepancies were verified by restriction enzyme analysis on two or more independent cDNA clones.

A comparison of the murine c-myb mRNA and published portions of the avian c-*myb* (16) gene are shown boxed in Fig. 2. The homologous region occurs at nucleotides 478–1599 (amino acid residues 71–444) in our murine c-myb mRNA sequence. There is 82% amino acid and 77% nucleotide sequence homology over this region. It is possible to identify three subregions that have even more striking homology than is apparent for the entire region by arbitarily dividing this 1122 nucleotide stretch into six subregions as follows: I (478–867); II (868–1041); III (1042–1227); IV (1228–1353); V (1354–1494); and VI (1495–1599). The amino acid homologies for these six subregions are 100, 45, 99, 67, 94, and 55%, respectively. In fact the minimal nonhomology in subregions III and V results from conservative amino acid substitutions in each case. Since the avian c-*myb* sequence was inferred by homology to the AMV v-*myb* sequence (16), both sequences are homologous to the same section of the murine c-*myb* sequence (nucleotides 478–1599). We note that the only nonconservative changes between the avian c-*myb* and v-*myb* are four amino acid differences spread throughout subregion I. The E26 virus v-*myb* sequence (17) is homologous to nucleotides 506–1350 (amino acid residues 81–361) in the mouse c-*myb* sequence and thus lacks the first 10 amino acids in subregion I and all of subregions V and VI. It was initially reported that avian c-*myb* and v-*myb* contain two tandem repeats of approximately 50 amino acids each in subregion I (33). Gonda *et al.* (29) have pointed out that their murine cDNA sequence and the published *Drosophila* genomic sequence (34) include three tandem repeats of amino acids (corresponding to nucleotides 377–532, 533–688, and 689–841 in our sequence). In addition to this intriguing observation regarding tandem repeats in subregion I, the remarkable amino acid sequence conservation of subregions I, III, and V suggests that there may be multiple noncontiguous regions in the c-myb protein that are important for its functions.

All of the apparent heterogeneity detected by S1 nuclease analysis of c-myb mRNA (Figs. 4 and 5) occurs to the 5′ side

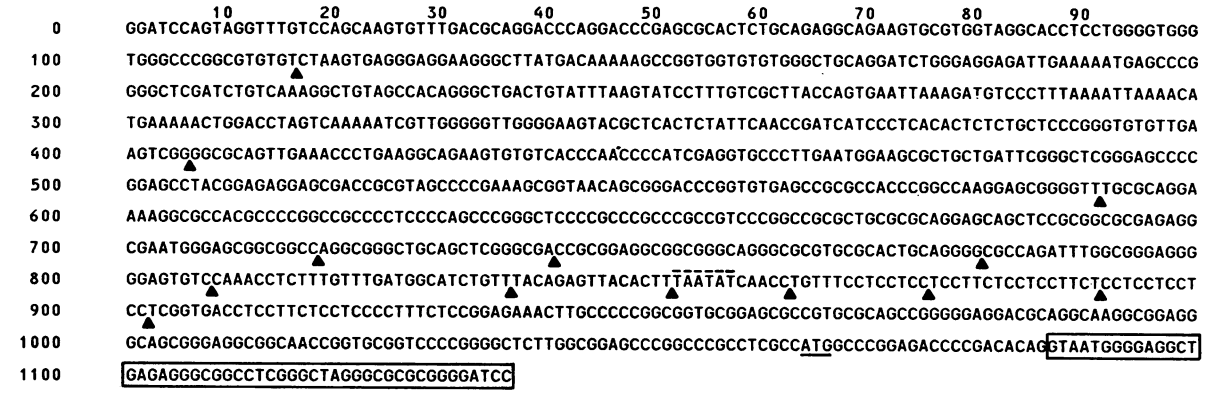

FIG. 4. Nucleotide sequence and organization of the 1136-bp *Bam*HI murine c-*myb* genomic fragment containing the translation initiation codon. The sequence includes bases 1–287 of the composite mRNA at positions 800–1086. The translation initiation codon is underlined. The boxed region contains 50 bp of intron sequence that begins at a consensus splice donor site. This sequence also includes 799 bp of putative 5′-untranslated region and flanking sequence. The 5′ extremities of S1 nuclease resistant products seen in Fig. 5 are marked (▲). A potential TATA sequence is marked by dashes.

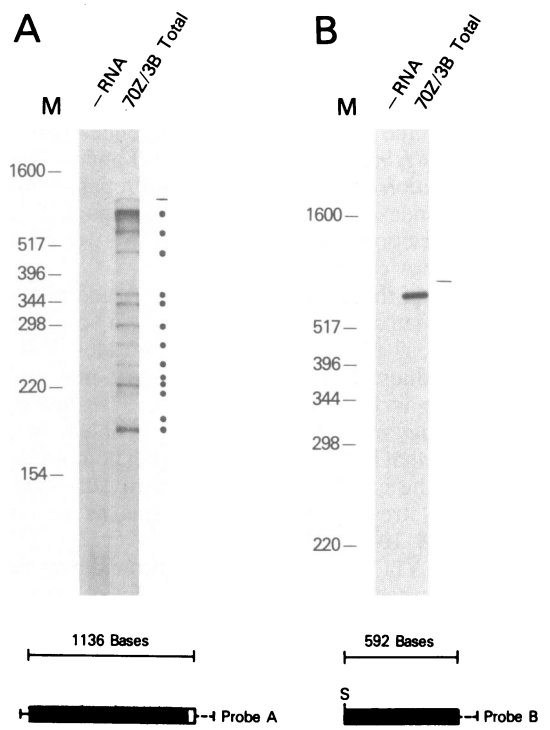

FIG. 5. S1 nuclease protection assay by 70Z/3B total cellular RNA of single-stranded DNA probes prepared by primer extension of M13 clones. (*A*) Probe A is the 1136-base *Bam*HI genomic fragment described in Fig. 4. It contains bases 1–287 of our c-myb mRNA sequence and 799 bases of putative 5′-untranslated region and flanking sequence (solid portion) plus 50 bases of intron (open portion). As indicated by thin lines the probe carries 22 bases of M13-polylinker DNA on the 3′ side of the genomic fragment and 46 bases at its 5′ end including the 17-base primer. Specific S1 nuclease resistant products are marked to the right with dots. A band representing undigested probe is marked with a dash. (*B*) Probe B contains bases 233–824 of the cDNA sequence (solid region) plus the 17-base primer at its 5′ end (thin line). Molecular size markers (M) are *Hin*fI restricted pBR322 in bases.

of the proposed translation initiation codon and thus represents mRNA capable of being translated into a single protein. However, the mechanism and biological relevance of this extreme heterogeneity is not readily apparent though 5′ heterogeneity has been reported in viral and eukaryotic cellular genes (35–39). Some of these examples (35–37) have been shown to be due to multiple transcription initiation sites, with transcripts generally initiating in association with classical promoter structures. Though there is a single TATA type of structure at position 852 (Fig. 4), no CAAT box is associated with it. The promoter regions of an increasing number of cellular genes, as well as of simian virus 40, have been shown to be very G+C-rich and to contain multiple copies of GC box sequences (i.e., CCGCCC or its inverted repeat) in lieu of or in addition to TATA and/or CAAT box structures (38, 39). We note that the 5′-untranslated/5′-flanking region of murine c-*myb* (Fig. 4) is very G+C rich (64%) and also contains three CCGCCC sequences (beginning at nucleotides 619, 643, and 647 in Fig. 4).

Our results on the genomic organization of the c-*myb* locus indicate that the 7.8-kb *Eco*RI genomic fragment is located on the 5′ side of the 4.2-kb *Eco*RI genomic fragment and contains the methionine codon at which translation is initiated (Fig. 4). Thus, the murine plasmacytoid lymphosarcoma tumors with Moloney murine leukemia virus insertions into the 4.2-kb *Eco*RI c-*myb* fragment (19) are disrupted in the c-*myb* coding region, so that the abnormal c-myb RNA products observed in these tumors probably lack the coding region for the normal amino terminus of the c-myb protein. The abnormality at the amino terminus of c-myb in these tumors would then be analogous to the deletion of the amino terminus of c-myb that exists for the v-myb proteins produced by both the AMV and E26 avian leukemia viruses (18).

We thank Dr. E. Sausville for help in preparing the cDNA libraries and Drs. J. Battey, G. Hollis, I. Kirsch, J. Minna, E. Sausville, and S. Segal for advice and critical reading of this manuscript. We also thank J. Boris for technical assistance. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Navy or the Department of Defense.

1. Boyle, W. J., Lampert, M. A., Lipsick, J. S. & Baluda, M. A. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4265–4269.
2. Klempnauer, K., Symonds, J., Evan, G. I. & Bishop, J. M. (1984) *Cell* **37,** 537–547.
3. Moelling, K., Pfaff, E., Beug, H., Beimling, P., Bunte, T., Schaller, H. E. & Graf, T. (1985) *Cell* **40,** 983–990.
4. Coll, J., Saule, S., Martin, P., Raes, M. B., Lagrou, C., Graf, T., Beug, H., Simon, I. E. & Stehelin, D. (1983) *Exp. Cell Res.* **149,** 151–162.
5. Mally, M. I., Vogt, M., Swift, S. E. & Haas, M. (1985) *Virology* **144,** 115–126.
6. Roy-Berman, P., Devi, B. G. & Parker, J. W. (1983) *Int. J. Cancer* **32,** 185–191.
7. Gonda, T. J., Sheiness, D. K. & Bishop, J. M. (1982) *Mol. Cell. Biol.* **2,** 617–624.
8. Sheiness, D. K. & Gardinier, M. (1984) *Mol. Cell. Biol.* **4,** 1206–1212.
9. Westin, E. H., Gallo, R. C., Arya, S. K., Eva, A., Souza, L. M., Baluda, M. A., Aaronson, S. A. & Wong-Staal, F. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 2194–2198.
10. Kirsch, I. R., Bertness, V., Silver, J. & Hollis, G. (1985) in *Leukemia: Recent Advances in Biology and Treatment,* UCLA Symposia on Molecular and Cellular Biology, eds. Gale, R. P. & Golde, D. W. (Liss, New York), Vol. 28, pp. 91–98.
11. Gonda, T. J. & Metcalf, D. (1984) *Nature (London)* **310,** 249–251.
12. Graf, T., Von Kirchbach, A. & Beug, H. (1981) *Exp. Cell Res.* **131,** 331–343.
13. Radke, K., Beug, H., Kornfield, S. & Graf, T. (1982) *Cell* **31,** 643–653.
14. Moscovici, M. G., Jurdic, P., Samarut, J., Gazzolo, L., Mura, C. V. & Moscovici, C. (1983) *Virology* **129,** 65–78.
15. Rushlow, K. E., Lautenberger, J. A., Papas, T. S., Baluda, M. A., Perbal, B., Chirikjian, J. G. & Reddy, E. P. (1982) *Science* **216,** 1421–1423.
16. Klempnauer, K., Gonda, T. J. & Bishop, J. M. (1982) *Cell* **31,** 453–463.
17. Nunn, M., Weiher, H., Bullock, P. & Duesberg, P. (1984) *Virology* **139,** 330–339.
18. Klempnauer, H., Ramsay, G., Bishop, J. M., Moscovici, M. G., Moscovici, C., McGreth, J. P. & Levinson, A. D. (1983) *Cell* **33,** 345–355.
19. Shen-Ong, G. L., Potter, M., Mushinski, J. F., Lavu, S. & Reddy, E. P. (1984) *Science* **226,** 1077–1080.
20. Pelicci, P. G., Lanfrancone, L., Brathwaite, M. D., Wolman, S. R. & Dalla-Favera, R. (1984) *Science* **224,** 1117–1121.
21. Paige, C. J., Kincade, P. W. & Ralph, P. (1978) *J. Immunol.* **121,** 641–647.
22. Chirgwin, J., Aeybyle, A., McDonald, R. & Rutter, W. (1979) *Biochemistry* **18,** 5294–5299.
23. Shapiro, D. J. & Schimke, R. T. (1975) *J. Biol. Chem.* **250,** 1759–1764.
24. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2,** 161–170.
25. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
26. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1199.
27. Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G. & Leder, P. (1983) *Cell* **34,** 779–787.
28. Mushinski, J. F., Potter, M., Bauer, S. R. & Reddy, E. P. (1983) *Science* **220,** 795–798.
29. Gonda, T. J., Gough, N. M., Dunn, A. R. & de Blaquiere, J. (1985) *EMBO J.* **4,** 2003–2008.
30. Kozak, C. (1984) *Nucleic Acids Res.* **12,** 857–872.
31. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
32. Hieter, P. A., Hollis, G. F., Korsmeyer, S. J., Waldemann, T. A. & Leder, P. (1981) *Nature (London)* **294,** 536–540.
33. Ralston, R. & Bishop, J. M. (1983) *Nature (London)* **306,** 803–806.
34. Katzen, A. L., Kornberg, T. B. & Bishop, J. M. (1985) *Cell* **41,** 449–456.
35. Contreras, R., Gheysen, D., Knowland, J., van de Voorde, A. & Fiers, W. (1982) *Nature (London)* **300,** 500–505.
36. Allan, M., Lanyon, W. G. & Paul, J. (1983) *Cell* **35,** 187–197.
37. Selvanayagam, C. S., Tsai, S. Y., Tsai, M. J., Selvanayagam, P. & Saunders, G. F. (1984) *J. Biol. Chem.* **259,** 14642–14646.
38. Reynolds, A. G., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Goldstein, J. L. & Luskey, K. L. (1984) *Cell* **38,** 275–285.
39. Ishii, S., Merlino, G. T. & Pastan, I. (1985) *Science* **230,** 1378–1381.