Studies of the Human c-*myb* Gene and Its Product in Human Acute Leukemias

DENNIS J. SLAMON, THOMAS C. BOONE, DOUGLAS C. MURDOCK, DUANE E. KEITH, MICHAEL F. PRESS, RICHARD A. LARSON, LAWRENCE M. SOUZA

The *myb* gene is the transforming oncogene of the avian myeloblastosis virus (AMV); its normal cellular homolog, c-*myb*, is conserved across a broad span of evolution. In humans, c-*myb* is expressed in malignant hematopoietic cell lines and in primary hematopoietic tumors. Partial complementary DNA clones were generated from blast cells of patients with acute myelogenous leukemia. The sequences of the clones were compared to the c-*myb* of other species, as well as the v-*myb* of AMV. In addition, the carboxyl terminal region of human c-*myb* was placed in an expression vector to obtain protein for the generation of antiserum, which was used to identify the human c-*myb* gene product. Like v-*myb*, this protein was found within the nucleus of leukemic cells where it was associated with the nuclear matrix. These studies provide further evidence that c-*myb* might be involved in human leukemia.

AVIAN MYELOBLASTOSIS IS A LEUKEMIC process involving malignant transformation of cells of the granulocyte-monocyte lineage in chickens. The etiologic agent for the disease is the avian myeloblastosis virus (AMV) (*1*). The v-*myb* oncogene of AMV is directly responsible for induction and/or maintenance of leukemia in avian myeloblastosis (*2*). As with other viral oncogenes, v-*myb* derived from a normal cellular homolog (c-*myb*) (*3*), which has been identified in a large number of vertebrate species including man (*4*) and in an invertebrate species, *Drosophila melanogaster* (*5*). In the case of v-*myb*, a protein of 45 kilodaltons (kD) has been found in leukemic cells from AMV-infected chickens, while the normal chicken c-*myb* protein has been identified as a 75-kD protein (*6*). Studies on expression of the human c-*myb* gene have been limited to messenger RNA (mRNA). These studies have shown an association between expression of the gene and human hematopoietic malignancies both in vitro and in vivo (*7, 8*). This association indicates that the gene may be involved in the leukemic process and is underscored by the direct etiologic role of v-*myb* in myelomonocytic leukemias of chickens.

By means of a v-*myb* probe, we examined mRNA from circulating blasts of three patients with acute myelogenous leukemia (AML) and two with acute lymphoblastic leukemia (ALL), as well as mRNA from two human tumors that do not contain c-*myb* message. In blasts from all five patients with leukemia, a transcript of 3.4 kilobases (kb) was identified (Fig. 1A, lanes 1 to 5). Others have placed the size of the human c-*myb* transcript from leukemic cells at 4.5 kb (*9, 10*). Our estimate of the size was determined relative to ribosomal RNA, as well as to standards of known size that would hybridize with the probe—RNA from AMV-infected cells that contain the 7.2-kb genomic viral RNA and the 2.0-kb subgenomic v-*myb* transcript, as well as RNA from chicken thymocytes that express chicken c-*myb* transcripts (Fig. 1B, lanes 1 to 5). The 3.4-kb size of the transcript is more consistent with a 3.8-kb human c-*myb* transcript previously reported (*11*). To further establish that we were looking at the c-*myb* transcript in human cells, we used polyadenylated [poly(A)$^+$] RNA from the undifferentiated as well as the DMSO-differentiated HL-60 cell line. Our data confirm previous studies (*9*) that show that the 3.4-kb transcript is found in the undifferentiated HL-60 cells and absent in the differentiated cells (Fig. 1B, lanes 2 and 3).

Since the leukemic cells from the patient whose RNA is shown in lane 2 of Fig. 1A appeared to have the greatest amount of human c-*myb* transcripts, this mRNA and another containing similar levels of *myb* transcripts were selected for generation of a complementary DNA (cDNA) library. The library was constructed with the Okayama and Berg vectors (*12*). This technique ensures that all cDNA inserts are in a defined orientation and begin with poly(A)$^+$ tracts from a fixed point in the vector. The cDNA libraries contained approximately 65,000 and 400,000 clones, respectively. Both banks were screened with a v-*myb* probe as described in Fig. 2, and 21 human c-*myb* cDNA clones were obtained. One clone was obtained from the library of 65,000 clones, and 20 clones came from the 400,000-clone library.

Restriction endonuclease mapping of the isolated clones defined two related but distinct classes of cDNA's (Fig. 2A). One class, represented by h-*myb*1 (six clones), contained cDNA inserts that were approximately 1000 base pairs (bp) or less in length. The clone obtained from the smaller library was

D. J. Slamon and D. E. Keith, Division of Hematology-Oncology, Department of Medicine, UCLA School of Medicine, Los Angeles, CA 90024.
T. C. Boone, D. C. Murdock, L. M. Souza, Amgen, Thousand Oaks, CA 91320.
M. F. Press and R. A. Larson, Departments of Pathology and Medicine, University of Chicago School of Medicine, Chicago, IL 60637.

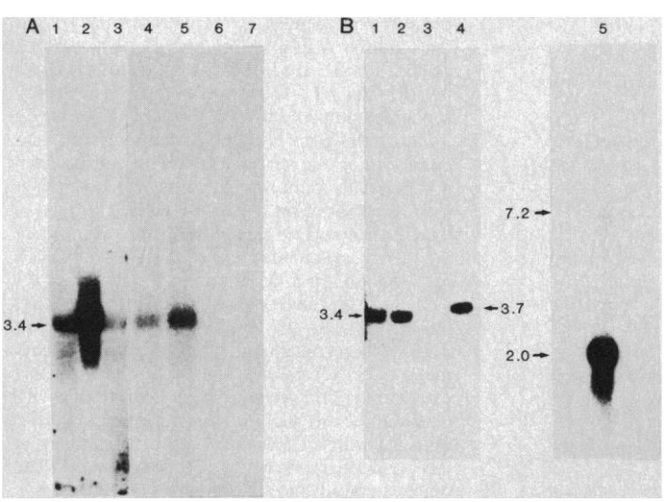
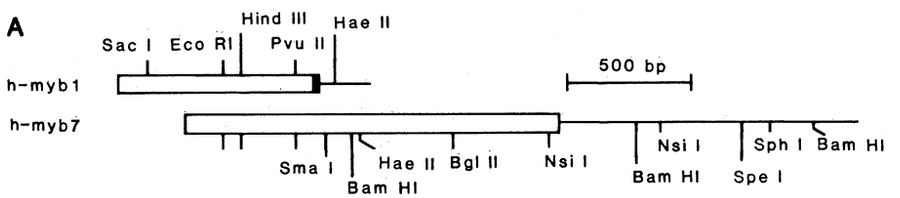


Fig. 1. (A) Northern blot analysis of poly(A)$^+$ RNA from leukemic blasts of patients with AML and ALL; (lanes 1 to 3) RNA from leukemic blasts of three patients with AML; (lanes 4 and 5) RNA from leukemic blasts of two patients with ALL; (lanes 6 and 7) RNA from human lung adenocarcinoma and colon adenocarcinoma tumor tissue that lack c-*myb* transcripts (*8*). (B) Northern blot analyses of poly(A)$^+$ RNA from leukemic blasts of a patient with AML and control cell lines. (Lane 1) RNA from leukemic blasts of the patient with ALL shown in (A), lane 5; (lane 2) RNA from the HL-60 promyelocytic human leukemia cell line; (lane 3) RNA from HL-60 cells induced to undergo differentiation to mature granulocytes by DMSO (*9*); (lane 4) RNA from normal chicken thymocytes that express chicken c-*myb* (*18*); (lane 5) RNA from the BM-2 avian myeloblast cell line, which has been transformed by AMV and expresses v-*myb* transcripts (*6*). RNA for all samples was prepared as previously described (*9*). Northern analysis (*27*) was performed on 10 mg of poly(A)$^+$ RNA in all cases.

a member of this class. The second class, represented by h-*myb*7 (15 clones), had cDNA inserts up to 2800 bp in length. Both h-*myb*1 and h-*myb*7 contained the largest inserts for their respective class. The sequence of h-*myb*1 diverged completely from h-*myb*7 at nucleotide position 811 (Fig. 2, B and C), and the h-*myb*1 open reading frame closed seven amino acids from this position (Fig. 2C). Comparison of the sequence of h-*myb*1 and h-*myb*7 with the genomic sequence of chicken c-*myb* revealed that the abrupt loss of homology between h-*myb*1 and h-*myb*7 occurs at the 3′ boundary of exon 5, relative to the chicken sequence.

An oligonucleotide probe made to span the h-*myb*1 unique fifth-sixth exon junction and its 3′ untranslated sequence [5′–TGATGACAGTCTTACTGGTAGCA-CCTGCTG–3′] was used to demonstrate the presence of these sequences in the other five clones classified with h-*myb*1 (*13*). This indicated that all six clones were truncated at the same point and had similar sequences beyond the fifth exon. Moreover, the inability of this oligonucleotide probe to hybridize with either plasmid or bacterial DNA (*13*) argues against a cloning artifact. The fact that the same truncated clones were isolated from libraries representing leukemic cell populations from two separate patients proves that these clones are not unique to an individual patient. Thus, the most likely explanation for the generation of truncated c-*myb* cDNA clones is the existence of alternatively spliced c-*myb* transcripts in humans.

Alternative splicing of the human c-*myb* gene is supported by the following: (i) the divergence of homology at a putative exon boundary, (ii) the presence of a polyadenylation signal at the end of both classes of *myb* clones, and (iii) the presence of altered c-*myb* transcripts in other species (*14*). The ratio of long to truncated c-*myb* clones was 3:1, yet we were unable to detect the presence of a second smaller c-*myb* message by Northern blot analysis with either a v-*myb* probe (Fig. 1) or an oligonucleotide probe specific for the unique human 3′ region [5′–CTATCCCCTCATTCAAGCACATGATG-ACAG–3′] in the truncated clones. This discrepancy may be explained by an unexpectedly low cDNA cloning efficiency of the 3.4-kb mRNA, since we would have anticipated at least a 30-fold increase in the number of nontruncated clones obtained, given the level of the 3.4-kb transcript pres-



B

GAG GAT GAA AAA CTG AAG AAG CTG GTG GAA CAG AAT GGA ACA GAT GAC TGG AAA GTT ATT GCC AAT TAT CTC CCG AAT CGA ACA GAT GTG
10 20 30 40 50 60 70 80 90

CAG TGC CAG CAC CGA TGG CAG AAA GTA CTA AAC CCT GAG CTC ATC AAG GGT CCT TGG ACC AAA GAA GAA GAT CAG AGA GTG ATA GAG CTT
100 110 120 130 140 150 160 170 180

GTA CAG AAA TAC GGT CCG AAA CGT TGG TCT GTT ATT GCC AAG CAC TTA AAG GGG AGA ATT GGA AAA CAA TGT AGG GAG AGG TGG CAT AAC
190 200 210 220 230 240 250 260 270

CAC TTG AAT CCA GAA GTT AAG AAA ACC TCC TGG ACA GAA GAG GAA GAC AGA ATT ATT TAC CAG GCA CAC AAG AGA CTG GGG AAC AGA TGG
280 290 300 310 320 330 340 350 360

GCA GAA ATC GCA AAG CTA CTG CCT GGA CGA ACT GAT AAT GCT ATC AAG AAC CAC TGG AAT TCT ACA ATG CGT CGG AAG GTC GAA CAG GAA
370 380 390 400 410 420 430 440 450

GGT TAT CTG CAG GAG TCT TCA AAA GCC AGC CAG CCA GCA GTG GCC ACA AGC TTC CAG AAG AAC AGT CAT TTG ATG GGT TTT GCT CAG GCT
460 470 480 490 500 510 520 530 540

CCG CCT ACA GCT CAA CTC CCT GCC ACT GGC CAG CCC ACT GTT AAC AAC GAC TAT TCC TAT TAC CAC ATT TCT GAA GCA CAA AAT GTC TCC
550 560 570 580 590 600 610 620 630

AGT CAT GTT CCA TAC CCT GTA GCG TTA CAT GTA AAT ATA GTC AAT GTC CCT CAG CCA GCT GCC GCA GCC ATT CAG AGA CAC TAT AAT GAT
640 650 660 670 680 690 700 710 720

GAA GAC CCT GAG AAG GAA AAG CGA ATA AAG GAA TTA GAA TTG CTC CTA ATG TCA ACC GAG AAT GAG CTA AAA GGA CAG CAG GTG CTA CCA
730 740 750 760 770 780 790 800 810

ACA CAG AAC CAC ACA TGC AGC TAC CCC GGG TGG CAC AGC ACC ACC ATT GCC GAC CAC ACC AGA CCT CAT GGA GAC AGT GCA CCT GTT TCC
820 830 840 850 860 870 880 890 900

TGT TTG GGA GAA CAC CAC TCC ACT CCA TCT CTG CCA GCG GAT CCT GGC TCC CTA CCT GAA GAA AGC GCC TCG CCA GCA AGG TGC ATG ATC
910 920 930 940 950 960 970 980 990

GTC CAC CAG GGC ACC ATT CTG GAT AAT GTT AAG AAC CTC TTA GAA TTT GCA GAA ACA CTC CAA TTT ATA GAT TCT TTC TTA AAC ACT TCC
1000 1010 1020 1030 1040 1050 1060 1070 1080

AGT AAC CAT GAA AAC TCA GAC TTG GAA ATG CCT TCT TTA ACT TCC ACC CCC CTC ATT GGT CAC AAA TTG ACT GTT ACA ACA CCA TTT CAT
1090 1100 1110 1120 1130 1140 1150 1160 1170

AGA GAC CAG ACT GTG AAA ACT CAA AAG GAA AAT ACT GTT TTT AGA ACC CCA GCT ATC AAA AGG TCA ATC TTA GAA AGC TCT CCA AGA ACT
1180 1190 1200 1210 1220 1230 1240 1250 1260

CCT ACA CCA TTC AAA CAT GCA CTT GCA GCT CAA GAA ATT AAA TAC GGT CCC CTG AAG ATG CTA CCT CAG ACA CCC TCT CAT CTA GTA GAA
1270 1280 1290 1300 1310 1320 1330 1340 1350

GAT CTG CAG GAT GTG ATC AAA CAG GAA TCT GAT GAA TCT GGA ATT GTT GCT GAG TTT CAA GAA AAT GGA CCA CCC TTA CTG AAG AAA ATC
1360 1370 1380 1390 1400 1410 1420 1430 1440

AAA CAA GAG GTG GAA TCT CCA ACT GAT AAA TCA GGA AAC TTC TTC TGC TCA CAC CAC TGG GAA GGG GAC AGT CTG AAT ACC CAA CTG TTC
1450 1460 1470 1480 1490 1500 1510 1520 1530

ACG CAG ACC TCG CCT GTG GCA GAT GCA CCG AAT ATT CTT ACA AGC TCC GTT TTA ATG GCA CCA GCA TCA GAA GAT GAA GAC AAT GTT CTC
1540 1550 1560 1570 1580 1590 1600 1610 1620

AAA GCA TTT ACA GTA CCT AAA AAC AGG TCC CTG GCG AGC CCC TTG CAG CCT TGT AGC AGT ACC TGG GAA CCT GCA TCC TGT GGA AAG ATG
1630 1640 1650 1660 1670 1680 1690 1700 1710

GAG GAG CAG ATG ACA TCT TCC AGT CAA GCT CGT AAA TAC GTG AAT GCA TTC TCA GCC CGG ACG CTG GTC ATG TGA GACATTTCCAGAAAAGCAT
1720 1730 1740 1750 1760 1770 1780

TATGGTTTTCAGAACACTTCAAGTTGACTTGGGATATATCATTCCTCAACATGAAACTTTTCATGAATGGGAGAAGAACCTATTTTTGTTGTGGTACAACAGTTGAGAGCAGCACCAAG

TGCATTTAGTTGAATGAAGTCTTCTTGGATTTCACCCAACTAAAAGGATTTTTAAAAATAAATAACAGTCTTACCTAAATTATTAGGTAATGAATTGTAGCCAGTTGTTAATATCTTAA

TGCAGATTTTTTTAAAAAAAAACATAAAATGATTTATCTGTATTTTAAAGGATCCAACAGATCAGTATTTTTTCCTGTGATGGGTTTTTTGAAATTTGACACATTAAAAGGTACTCCAG

TATTTCACTTTTCTCGATCACTAAACATATGCATATATTTTTAAAAATCAGTAAAAGCATTACTCTAAGTGTAGACTTAATACCATGTGACATTTAATCCAGATTGTAAATGCTCATTT

ATGGTTAATGACATTGAAGGTACATTTATTGTACCAAACCATTTTATGAGTTTTCTGTTAGCTTGCTTTAAAAATTATTACTGTAAGAAATAGTTTTATAAAAAATTATATTTTTTATTC

AGTAATTTAATTTTGTAAATGCCAAATGAAAAACGTTTTTTGCTGCTATGGTCTTAGCCTGTAGACATGCTGCTAGTATCAGAGGGGCAGTAGAGCTTGGACAGAAAGAAAAGAAACTT

GGTGTTAGGTAATTGACTATGCACTAGTATTTCAGACTTTTTAATTTTATATATATATACATTTTTTTCCTTCTGCAATACATTTGAAAACTTGTTTGGGAGACTCTGCATTTTTTAT

TGTGGTTTTTTTGTTATTGTTGGTTTATACAAGCATGCGTTGCACTTCTTTTTTGGGAGATGTGTGTTGTTGATGTTCTATGTTTTGTTTTGAGTGTAGCCTGACTGTTTTATAATTTG

GGAGTTCTGCATTTGATCCGCATCCCCTGTGGTTTCTAAGTGTATGGTCTCAGAACTGTTGCATGGATCCTGTGTTTGCAACTGGGGAGACAGAAACTGTGGTTGATAGCCAGTCACTG

CCTTAAGAACATTTGATGCAAGATGGCCAGCACTGAACTTTTGAGATATGACGGTGTACTTACTGCCTTGTAGCAAAATAAAGATGTGCCCTTATTTTACCTpA

C

GTA AGA CTG TCA TCA TGT GCT TGA ATGAGGGGATAGCAGCTTTGCCTCAGTTTACCTAAGCGCTCTTCTCTTCTAAATATTACACTTAGCAAGGCTCCATATATCCATTCA

GAATGTCTCAACACAAGAAGTTGCTTGTAGTAAAATGTAGTTGGTATCAGATTATATGCTGATTAAATTGGAAGCAGTCTTTTTGTAATTGCAATAAAAATGCAATATCCACTTpA

Fig. 2. (A) Restriction endonuclease map of cDNA clones h-*myb*1 and h-*myb*7. The cDNA inserts are aligned with each other on the basis of their common enzyme sites and shared sequences. The open reading frame in each clone is depicted by an open bar with the enclosed portion of the bar for h-*myb*1, representing the coding region unique to the truncated cDNA clones. The line extending from each bar represents the 3′ untranslated sequence. (B) A composite nucleotide sequence of human c-*myb* obtained from cDNA clones h-*myb*1 and h-*myb*7. An arrow opposite nucleotide 810 designates the point at which the sequence in h-*myb*1 diverges from h-*myb*7. The putative polyadenylation signal is underlined and the coding region is numbered. The cDNA clones, h-*myb*1 and h-*myb*7, were identified by an [α-^{32}P]–labeled v-*myb* probe (*28*). The v-*myb* probe was prepared from plasmid HAX-4 (*29*) by isolating the Bam HI–Xba I fragment and nick-translating it to a specific activity of approximately 1.5×10^8 count/min per microgram of DNA (*30*). Hybridization of the ^{32}P-labeled probes to the colony-containing filters and subsequent washing was carried out as described (*31*). Plasmid DNA for subcloning into M13 vectors mp10 and mp11 (*32*) was prepared as described (*33*). Dideoxy sequencing (*34*) was performed on independent M13 clones containing either overlapping fragments or cDNA strands. Sequences found in brackets represent the conserved tandem repeats discussed in the text (*15*). (C) The nucleotide sequence obtained from cDNA clone h-*myb*1 beginning at the point of divergence (811) from h-*myb*7. The putative polyadenylation signal is underlined.

ent in the two mRNA preparations used for cloning.

The composite c-*myb* nucleotide sequence (Fig. 2B) represents the 3′ 2977 bases of the 3.4-kb message. The largest open reading frame in the c-*myb* sequence has the potential to encode the COOH-terminal 594 amino acids of the human c-*myb* protein (Fig. 2B). We then compared the deduced protein sequences for human, chicken, and mouse c-*myb* based on alignment with the seven known exons for chicken c-*myb* (Fig. 3). The overall sequence homology found when comparing the human c-*myb* to chicken and mouse c-*myb* was 85% and 90%, respectively. There was 100% homology of amino acids between chicken and human c-*myb* in exons 1 and 2, with only one amino acid difference in this region of mouse protein (*15*). Moreover, the three repeat sequences in this region of the mouse protein were also present in the human protein. Two of these repeats have been identified in the chicken (*16*) and *Drosophila* c-*myb* genes (*5*), and it has been postulated that these evolutionarily conserved tandem repeats are involved in *myb* protein–DNA interactions (*15*). The human c-*myb* diverged furthest from the chicken c-*myb* in exons 3, 4, 6, and 7, with 68%, 85%, 81%, and 77% homology, respectively (Fig. 3). Like exons 1 and 2, the fifth exon exhibited a high degree (97%) of homology between the avian and human proteins. The same relative percentages of exon homology were seen when comparing the c-*myb* protein sequence from human to mouse and mouse to chicken, with the exception of the area encoded by exon 6 (Fig. 3). From the deduced amino acid sequence, we would expect the human c-*myb* protein to be larger than 65 kD.

Construction of an expression vector was undertaken to generate leukemic human c-*myb* antigen in *Escherichia coli* for the purpose of raising polyvalent antibodies. The construct was made by fusing, in frame, the 5′ region of a bovine growth hormone (bGH) gene to the region containing exons 1–5 of human c-*myb* gene (Fig. 4). This linkage would produce a fusion protein containing the NH_2-terminal 76 amino acids of bGH and 228 amino acids of human c-*myb*. Whole bacterial cell extracts containing this construct yielded the expected inducible protein of 38 kD when analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). The fusion protein was purified from whole-cell bacterial lysates by electrophoresis (SDS-PAGE) (*17*) for injection into rabbits.

Polyclonal antisera to the bGH/human c-*myb* fusion protein were generated as described (Fig. 5) and were characterized by their ability to recognize *myb* gene products.

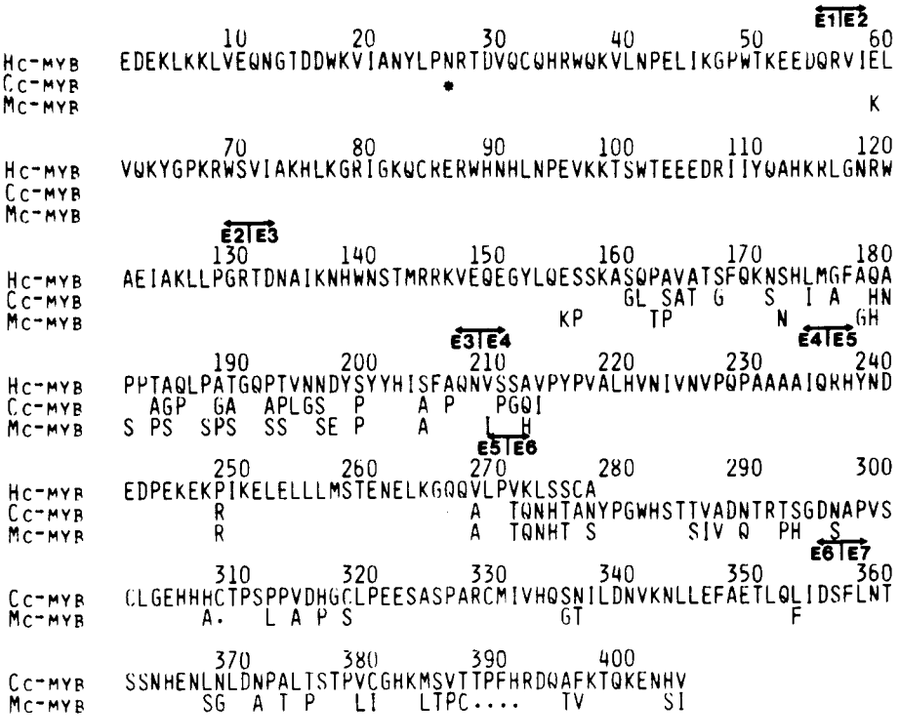


Fig. 3. The deduced amino acid sequence derived from the human nucleotide sequence of cDNA clone aligned with both the chicken (*7*) and mouse c-*myb* (*15*) sequences. Only those amino acids in the mouse and chicken c-*myb* sequences differing from the human are depicted. The comparison with chicken c-*myb* begins at position 26 and ends at position 404 (denoted by an asterisk). The exon boundaries for the chicken c-*myb* sequence are displayed by arrows above the single-letter amino acid code (*35*). The dots in the mouse and human sequences represent shifts in the alignment of these sequences relative to the chicken to maintain homology.

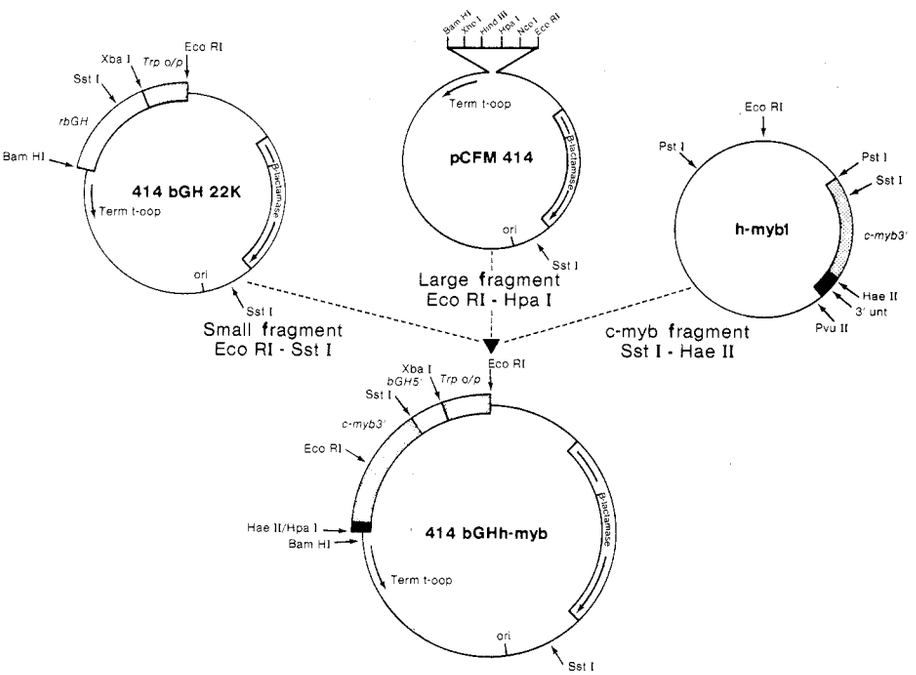


Fig. 4. Expression of a bGH/human c-*myb* fusion protein in *E. coli*. The expression vector was constructed through a three-way ligation involving restriction endonuclease fragments from three separate plasmids. The final construct consists of a modified ts-runaway plasmid (*36*), a tryptophan synthetase promoter (*37*), and a fusion gene between the NH_2-terminal portion of bGH and the COOH-terminal portion of human c-*myb*. Term t-oop is a λ phage transcription terminator. The c-*myb* fragment was generated by digestion with Hae II followed by treatment with the Klenow fragment and digestion with Sst I.

Three areas of the deduced human c-*myb* protein sequence were nearly 100% homologous to the amino acids encoded by exons 1, 2, and 5 of the chicken c-*myb* gene and the analogous region in v-*myb* (Fig. 3). The polyclonal antiserum generated to the bGH/human c-*myb* fusion protein recognized avian *myb* proteins of appropriate sizes (Fig. 5A, lanes 1 and 2), which we predicted on the basis of the shared regions of amino acid sequence.

A protein of 83 kD was seen (Fig. 5A, lane 3) in leukemic blasts from the patient with AML whose RNA is shown in Fig. 1, lane 1. This protein was not seen when preimmune rabbit sera was used (Fig. 5A, lane 6), and recognition of the protein could be eliminated by pre-incubation of the antisera with the fusion protein (Fig. 5A, lane 4). Preincubation with bGH alone did not eliminate recognition of the 83-kD protein, indicating that precipitation of this protein was the result of antibodies directed against human c-*myb*–specific epitopes and not those of bGH (Fig. 5A, lane 5). In addition, the 83-kD protein was seen in three other patients with leukemia (two with AML and one with ALL), as well as in the HL-60 promyelocytic and MOLT-4 T-lymphocytic human leukemia cell lines (*13*).

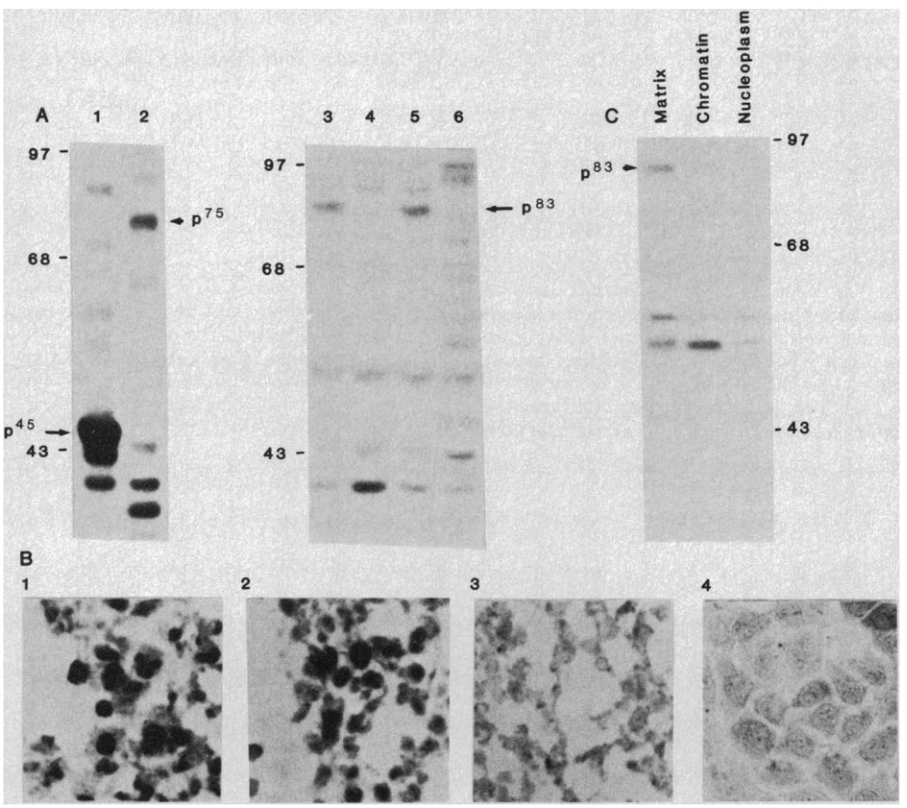


Fig. 5. (A) Immunoprecipitation of lysates from cells that express *myb* proteins. (Lane 1) BM-2 avian myeloblast cell lysate and antiserum to the bGH/human c-*myb* fusion protein. BM-2 cells express significant amounts of $p45^{v\text{-}myb}$ (p45) (*6*). (Lane 2) Chicken thymocyte cell lysate and antiserum to the fusion protein. Chicken thymocytes express significant amounts of $p75^{c\text{-}myb}$ (p75) (*18*). (Lane 3) Human leukemic myeloblast cell lysate and antiserum to the fusion protein from the patient represented in lane 1 of Fig. 1. (Lane 4) Human leukemic myeloblasts and antiserum to the fusion protein preincubated with the fusion protein. (Lane 5) Human leukemic myeloblasts and antiserum to the fusion protein preincubated with bGH alone. (Lane 6) Human leukemic myeloblasts and preimmune serum from the rabbit producing antiserum to the fusion protein; $p83^{c\text{-}myb}$ is shown as p83. Antibodies were raised in rabbits by injection with the gel-purified protein as previously described (*38*). All cells were metabolically labeled with [^{35}S]methionine, lysed, and subjected to immunoprecipitation with 10 μg of antisera (*38*). The other bands seen in lanes 3 to 6 were not competed away by the fusion protein; therefore, they were not the result of a specific immune reaction and represent background proteins seen with the immunoprecipitation reaction. (B) Immunocytochemical staining of cells containing *myb* gene product. (Panel 1) BM-2 avian myeloblasts and antiserum to the fusion protein; (panel 2) human leukemic myeloblasts from the patient represented in lane 1 of Fig. 1A and antiserum to the fusion protein; (panel 3) human leukemic myeloblasts and pre-immune serum; (panel 4) MCF-7 human breast cancer cells and antiserum to the fusion protein. Immunocytochemical staining was done by the immunoperoxidase method (*20*). (C) Subnuclear localization of the human c-*myb* protein. Nuclei from [^{35}S]methionine-labeled human leukemic myeloblasts were isolated and fractionated into a matrix, chromatin, and nucleoplasm fraction as previously described (*21*). The fractions were subjected to immunoprecipitation with 10 μg of anti-bGH/*myb* antisera. The fractions did not show evidence of cross-contamination when stained for matrix and histone proteins with Coomassie blue after electrophoresis on SDS-PAGE (*13, 21*).

The gene products of both v-*myb* and chicken c-*myb* are found in the nuclei of cells expressing these proteins (*18, 19*). To determine the cellular localization of the human c-*myb* protein, the antiserum to human *myb* was used in immunocytochemical staining of human leukemic cells. By means of an immunoperoxidase staining procedure (*20*), the antiserum to the bGH/human c-*myb* fusion protein localized the gene product of v-*myb* to the nuleus of AMV-transformed cells (Fig. 5B, panel 1). When leukemic blasts from a patient with acute myeloblastic leukemia were examined, the majority of the *myb*-specific staining was again seen in the nucleus of the cells (Fig. 5B, panel 2). There was no staining of human leukemic myeloblasts with preimmune sera (Fig. 5B, panel 3). Furthermore, the MCF-7 human breast cancer cell line (which does not contain detectable c-*myb* transcripts) was used as an additional control and no staining, nuclear or otherwise, occurred with the antisera to human *myb* (Fig. 5B, panel 4).

The majority of the nuclear v-*myb* protein in chicken myeloblasts is associated with the nuclear matrix (*21*). To determine if this is also true for the human c-*myb* protein, nuclei from human leukemic blasts were isolated and fractionated into nucleoplasm, chromatin, and matrix, as previously described (*21*). The separate fractions were then tested in an immunoprecipitation assay with the antiserum to bGH/human c-*myb*. The majority of the human p83 was found in the nuclear matrix fraction (Fig. 5C).

The evidence linking the human c-*myb* proto-oncogene to human leukemia is largely circumstantial and is derived from both in vitro and in vivo systems. First, human c-*myb* gene transcripts have been found almost exclusively in malignant hematopoietic cell lines in culture (*7*). A notable exception to this is the amplification and expression of the human c-*myb* gene in two independently derived cell lines from a human colon adenocarcinoma (*22*). Second, high levels of expression of the c-*myb* gene have been found in human leukemic blasts (KG-1) or leukemic promyelocytes (HL-60) in culture, but are undetectable when these cells are induced to undergo differentiation to either mature granulocytes (*9*) or macrophages (*23*). Finally, in a survey of 53 primary human tumors representing 20 different tumor types, the c-*myb* gene was found to be significantly expressed in only hematopoietic malignancies, with the highest levels occurring in myeloid and lymphoid leukemias (*8*).

The current study provides further circumstantial evidence linking the c-*myb* gene

to human leukemia. The data presented show that not only is there significant sequence homology between v-*myb* and human c-*myb*, but that within human leukemic cells, the $p83^{c\text{-}myb}$ protein behaves like the leukemogenic $p45^{v\text{-}myb}$, in that it is localized to the nucleus and is associated with the nuclear matrix. This nuclear localization may be of particular interest, considering recent experiments demonstrating that differentiation of avian myeloblasts to macrophages is accompanied by relocation of the v-*myb* protein from the nucleus to the cytoplasm (*19*). This has been taken to imply that nuclear localization is important in maintaining cell transformation (*24*).

Comparisons of quantitative and qualitative changes in the c-*myb* transcript and/or gene product in normal and malignant cells may give insights into the role of this gene in human leukemia. Examples of both types of changes in proto-oncogene expression can potentially lead to cell transformation (*25, 26*). Of note in the current study is the marked increase, relative to the other patients, of c-*myb* transcripts in the two patients with AML from whom the cDNA clones were derived. All six patients examined had 95 to 100% leukemic blasts in their peripheral blood. Identical amounts of poly(A)$^+$ RNA from each patient were analyzed by Northern blot analysis. However, unlike the other four patients who had between 45,000 to 80,000 cells per cubic millimeter, these two patients had greater than 205,000 cells per cubic millimeter. Three other patients with similar levels of c-*myb* transcripts have been identified, and all have had peripheral blood leukemic blast counts in excess of 190,000 cells per cubic millimeter. It is tempting to speculate that the level of c-*myb* transcripts may correlate with or be related to the peripheral blood blast count; however, a larger survey of several patients will be required to address this. Correlation of other disease parameters with c-*myb* expression such as prognosis and response to therapy may also be possible in such a study.

REFERENCES AND NOTES

1. M. A. Baluda and P. P. Jamieson, *Virology* **14**, 33 (1961); C. Moscovici, *Curr. Top. Microbiol. Immunol.* **71**, 79 (1975); D. Boettiger and E. Durban, *J. Virol.* **49**, 841 (1984).
2. M. G. Moscovici and C. Moscovici, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1421 (1983); M. G. Moscovici *et al.*, *Mol. Cell. Biol.* **5**, 3301 (1985).
3. L. M. Souza, M. C. Komaromy, M. A. Baluda, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3004 (1980); T. J. Gonda, *Cell* **23**, 279 (1981); K. E. Rushlow *et al.*, *J. Virol.* **42**, 840 (1982); L. M. Souza *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5177 (1980); K.-H. Klempnauer, T. J. Gonda, J. M. Bishop, *Cell* **31**, 453 (1982).
4. D. G. Bergmann, L. M. Souza, M. A. Baluda, *J. Virol.* **40**, 450 (1981); D. Sheiness and J. M. Bishop, *ibid.* **31**, 514 (1979).
5. A. L. Katzen, T. B. Kornberg, J. M. Bishop, *Cell* **41**, 449 (1985).
6. W. J. Boyle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2834 (1983); K.-H. Klempnauer *et al.*, *Cell* **33**, 345 (1983).
7. E. H. Westin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2490 (1982).
8. D. J. Slamon *et al.*, *Science* **224**, 256 (1984).
9. E. H. Westin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2194 (1982).
10. M. Blick *et al.*, *Blood* **64**, 1234 (1984).
11. D. LePrince *et al.*, *EMBO J.* **2**, 1073 (1982).
12. H. Okayama and P. Berg, *Mol. Cell. Biol.* **2**, 161 (1982).
13. D. J. Slamon *et al.*, unpublished data.
14. G. L. Shen-Ong *et al.*, *Science* **226**, 1077 (1984).
15. T. J. Gonda *et al.*, *EMBO J.* **4**, 2003 (1985).
16. R. Ralston and J. M. Bishop, *Nature (London)* **306**, 803 (1983).
17. W. M. Hunkapiller *et al.*, *Methods Enzymol.* **91**, 227 (1983).
18. W. J. Boyle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4265 (1984).
19. K.-H. Klempnauer *et al.*, *Cell* **37**, 537 (1984).
20. L. A. Sternberger, in *Immunocytochemistry* (Wiley, New York, 1978), pp. 104–169.
21. W. J. Boyle *et al.*, *Mol. Cell. Biol.* **5**, 3017 (1985).
22. K. Alitalo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4534 (1984).
23. D. J. Slamon and H. P. Koeffler, in preparation.
24. J. M. Bishop, *Cell* **42**, 23 (1985).
25. J. Erikson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 820 (1983).
26. C. J. Tabin *et al.*, *Nature (London)* **300**, 143 (1982).
27. P. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5201 (1980).
28. W. D. Benton and R. W. Davis, *Science* **196**, 180 (1977).
29. B. Perbal and M. A. Baluda, *J. Virol.* **41**, 250 (1982).
30. T. Manniatis *et al.*, *Cell* **8**, 163 (1976).
31. G. M. Wahl, M. Stern, G. R. Stark, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3683 (1979).
32. J. Messing, *Methods Enzymol.* **101**, 20 (1983).
33. R. D. Klein, E. Selsing, R. D. Wells, *Plasmid* **3**, 88 (1980).
34. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
35. The following abbreviations were used for amino acids: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.
36. B. E. Uhlin and K. Nordstrom, *Mol. Gen. Genet.* **165**, 167 (1978).
37. G. F. Miozzari and C. Yanofsky, *J. Bacteriol.* **133**, 1457 (1978).
38. D. J. Slamon *et al.*, *Science* **226**, 61 (1984).

39. Supported by U.S. Public Health Service grant CA 36827. We thank W. J. Boyle of UCLA for providing the BM-2 cell line, M. A. Baluda for providing the HAX-4 v-*myb* fragment, and H. P. Koeffler for providing undifferentiated and DMSO-differentiated HL-60 cells. We also thank L. Ramos, R. Cortini, N. Nousek-Goebl, and J. Udove for their technical assistance.

11 November 1985; accepted 26 March 1986

Structure, Antigenic Determinants of Some Clinically Important Insect Allergens: Chironomid Hemoglobins

XAVER BAUR, HEINZ ASCHAUER, GERTRAUD MAZUR, MAHMOUD DEWAIR, HUGO PRELICZ, WOLFGANG STEIGEMANN

Determination of the molecular structure and properties of allergens that elicit severe immediate-type hypersensitivity diseases in humans and a knowledge of the structure of their antibody-binding sites should provide new insight into the pathogenetic mechanisms of allergic diseases. Monomeric and homodimeric hemoglobins (CTT I to X) have been identified as potent allergenic components of Chironomidae, a family of Diptera. Immunologic investigations of peptides of three of these hemoglobins (CTT IV, CTT VI, and CTT VIII) showed that human antibodies of the E and G classes recognize at least two different sites within each molecule. Individual hemoglobin peptides were aligned with homologous regions of chironomid hemoglobin CTT III, whose tertiary structure has been determined by x-ray analysis at a resolution of 1.4 angstroms. The antigenic site CTT IV(91 to 101) showed the following characteristics: (i) seven polar or hydroxylated amino acids, from a total of eleven, occupying predominantly superficial regions; (ii) the property of linkage to other molecules by hydrogen bonds or solvent clusters; and (iii) high thermal mobility factors. In contrast, peptide CTT IV(102 to 108), which does not bind human antibodies, contained no polar amino acids and had low thermal mobility factors. These results support the idea that the antigenicity of clinically relevant proteins is related to regions with a predominance of polar amino acids and with low energy barriers between different conformations, which allow high flexibility, including site-specific adaptation in antibody binding.

CONTACT WITH BOTH LARVAE AND midges of chironomids, nonbiting insects, elicits severe and immediate hypersensitivity in approximately 20% of environmentally or occupationally exposed persons (*1–3*).

A crude extract of larvae from one European chironomid species, *Chironomus thummi*, was fractionated (*1, 4–8*), and the fractions were studied by the radioallergosorbent test used to estimate allergen-specific antibodies of the immunoglobulin E class

Pneumologische Abteilung, Medizinische Klinik I, Klinikum Grosshadern, Universität München, München, Federal Republic of Germany; Institut und Poliklinik für Arbeitsmedizin, Universität München, München, Federal Republic of Germany; Max-Planck-Institut für Biochemie, Am Klopferspitz, 8033 Martinsried, Federal Republic of Germany.