Fetal and Variant α-Fetoproteins Are Encoded by mRNAs That Differ in Sequence at the 5′ End†

Yu-Jui Yvonne Wan, Juan L. Jimenez-Molina, and Janice Yang Chou*

Human Genetics Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892

Received March 1, 1988; Revised Manuscript Received May 31, 1988

ABSTRACT: The temperature-sensitive RLA209-15 fetal rat hepatocyte line grown at the nonpermissive temperature (40 °C, normal phenotype) produces authentic rat α-fetoproteins (AFPs) of 69K and 73K (fetal AFPs) which are encoded by a 2.2-kb mRNA. These cells also produce low levels of a 1.7-kb AFP mRNA and a 65K variant AFP when grown at the permissive temperature (33 °C, transformed phenotype). Hybrid-selected translation demonstrates that the 1.7-kb AFP mRNA encodes the 65K variant AFP. Northern blot hybridization and S1 nuclease analyses indicate that the 1.7-kb mRNA lacks sequences present in the first seven 5′ exons of the 2.2-kb AFP mRNA. However, the 1.7- and 2.2-kb AFP mRNAs share common sequences extending from the beginning of the eighth exon (corresponding to nucleotide 873 of the fetal AFP mRNA) to the 3′ end. Primer extension analysis suggests that the 1.7-kb RNA contains additional sequences 5′ to the common regions shared by both AFP mRNAs. We have previously shown that adult rat liver produces a 1.7-kb AFP mRNA; we now report the isolation of a cDNA (ARFP5) encoding this variant AFP mRNA from an adult rat liver cDNA library. Restriction endonuclease mapping and sequence analysis of ARFP5 confirm that the 1.7- and 2.2-kb AFP mRNAs share similar sequences at the 3′ region (approximately 1.1 kb). However, ARFP5 contains an additional 90 bp variant AFP mRNA-specific 5′ sequence which is located in the seventh intron of the rat AFP gene. In addition to RLA209-15 and adult rat liver, a rat hepatoma cell line (CRL1548) expresses the 1.7-kb but not the 2.2-kb AFP mRNA. The 1.7-kb AFP mRNA in CRL1548 cells, like the variant AFP mRNA of RLA209-15 cells, hybridizes to a probe containing either the 3′ sequence of the fetal AFP mRNA or the 90 bp 5′ sequence of ARFP5.

α-Fetoprotein (AFP) is the major serum protein of the fetus and is mainly produced by fetal liver and embryonic yolk sac (Abelev, 1971; Gitlin & Boesman, 1967). Production of AFP decreases to very low levels in adult life but reappears in adult serum in pathological conditions such as hepatocarcinomas and teratocarcinomeas (Abelev, 1974; Ruoslahti et al., 1974; Hirai, 1979). It has been suggested that AFP is the fetal counterpart of serum albumin and may play a role in the transport of estrogens (Abelev, 1971, 1974; Gitlin & Boesman, 1967). Although the exact physiological function of AFP has yet to be defined, expression of the AFP gene has provided a suitable model to study the molecular mechanisms responsible for tissue-specific gene regulation in differentiation, maturation, and carcinogenesis.

Only one AFP gene per haploid genome has so far been demonstrated in rat (Jagodzinski et al., 1981; Nahon & Sala-Trepat, 1984), mouse (Kioussis et al., 1981), and man (Sakai et al., 1985). In all three species, the AFP genes are organized similarly into 15 coding exons interrupted by 14 introns. The AFPs produced by fetal rat liver are two polypeptides of 69K and 73K (fetal AFPs) differing in carbohydrate side chains (Mano & Chou, 1982). The unglycosylated form of rat AFPs is a 66K polypeptide encoded by a mRNA of 2.2 kb (Jagodzinski et al., 1981). We have shown that RLA209-15 fetal rat liver cells produce a 1.7-kb AFP mRNA and a 65K variant AFP (Chou et al., 1982; Chou & Savitz, 1986). The RLA209-15 cell line was established by tranforming normal rat fetal hepatocytes with a temperature-sensitive (ts) A mutant of simian virus 40 (SV40). The SV40 A gene is required for maintenance of transformation (Chou & Schlegel-Haueter, 1981). At the nonpermissive temperature (40 °C), these cells are phenotypically nontransformed and produce mainly the 2.2-kb AFP mRNA and fetal AFP (Chou et al., 1982; Chou & Savitz, 1986). At the permissive temperature (33 °C), RLA209-15 cells are phenotypically transformed and produce low levels of the 2.2- and 1.7-kb AFP mRNAs and the fetal and variant AFPs. We proposed that the 1.7-kb AFP mRNA encodes the variant AFP because of their concurrent existence in RLA209-15 cells (Chou & Savitz, 1986). Since the 1.7-kb AFP mRNA binds strongly to cDNAs encoding the fetal AFP under stringent hybridization conditions, these two RNAs must share similar sequences.

In the present study, we show that the 1.7-kb AFP mRNA encodes the 65K AFP and lacks sequences present in the first seven 5′ exons of the 2.2-kb AFP mRNA but contains similar sequences to the fetal AFP mRNA extending from nucleotide 873 to the 3′ end of the fetal AFP mRNA. The existence of additional 5′ sequences specific for the 1.7-kb mRNA was demonstrated by the isolation and characterization of a cDNA clone encoding a major part of the 1.7-kb AFP mRNA.

Materials and Methods

Cell Culture. The SV40 tsA209 mutant virus-transformed fetal rat hepatocyte line RLA209-15 (Chou & Schlegel-Haueter, 1981) and the rat hepatoma cell line CRL1548 (American Type Culture Collection, Rockville, MD) were cultured at 33 and 37 °C, respectively, in α-modified minimal essential medium supplemented with streptomycin, penicillin, and 4% fetal bovine serum. RLA209-15 cultures maintained at 33 °C throughout growth were phenotypically transformed cells. RLA209-15 cultures grown initially at 33 °C, but shifted

† The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02838.

* Correspondence should be addressed to this author at NIH, Building 10, Room 8C429, Bethesda, MD 20892.

to 40 °C after 3–4 days of growth at 33 °C, were phenotypically differentiated cells.

Plasmids. RAF5PB, RAF5PS, and RAF5PP were constructed by subcloning the *Pst*I–*Bam*HI (nucleotides 1–720), *Pst*I–*Sst*I (nucleotides 1–954), and *Pst*I–*Pst*I (nucleotides 1–1529) fragments of PHDQ105 (containing nucleotides 1–1895 of the rat fetal AFP mRNA) (Jagodzinski et al., 1981; Turcotte et al., 1985) into pUC19, M13mp19, and pUC18, respectively. RAFPP and RAF3PP were constructed by subcloning the *Pst*I–*Pst*I (nucleotides 925–1529) fragment of pRAF65 (Jagodzinski, 1981) and the *Pst*I–*Pst*I (nucleotides 1529 to the 3′ end) fragment of pRAF6 (Jagodzinski, 1981) into M13mp18.

RAFGHA was constructed by subcloning the *Hinc*II–*Ava*I (198 bp, containing 126 bp of the Z exon with 55 bp of 5′ flanking and 17 bp of the first intron) fragment of rat AFP genomic clone pRAF65R0.5-7 (provided by Dr. J. Locker, University of Pittsburgh, Pittsburgh, PA) into Gemini 1 vector (Promega Biotec, Madison, WI). RAFGHH was constructed by subcloning the *Hin*dIII–*Hin*dIII (1.2 kb, located in the seventh intron of rat AFP gene) fragment of rat genomic clone PRAF7-E1 (J. Locker) into M13mp18 and M13mp19.

Hybrid-Selected Translation and Immunoprecipitation. Hybrid selection was performed according to the procedures described by Jacqueline et al. (1983). Single-stranded RAF3PP DNA (antisense, 40 μg/80 mm^2) was applied to nitrocellulose filters (Kafatos et al., 1979), baked 2 h at 80 °C, cut into 1 mm^2 pieces, and boiled in water for 1 min. Prehybridization was at 37 °C for 3 h in a buffer containing 50% formamide, 0.01 M Pipes (pH 6.4), 0.4 M NaCl, and 300 μg/mL calf liver tRNA. Hybridization was in the same buffer except the tRNA was replaced with 1 mg of total RNA isolated from RLA209-15 grown at 33 °C. After hybridization for 16 h, filters were washed 10 times with a buffer containing 150 mM NaCl, 15 mM sodium citrate, and 0.5% SDS for 1 min each at 60 °C and 3 times with 2 mM EDTA at 25 °C. RNA was eluted in 1 mM EDTA by boiling for 1 min at 100 °C and precipitated in the presence of 10 μg of calf liver tRNA. The selected RNA was dissolved in 10 μL of water, and 5 μL was used for in vitro translation.

A rabbit reticulocyte lysate system was employed for in vitro translation using L-[^{35}S]methionine as labeled precursor. AFPs in the translation products were isolated by immunoprecipitation with rabbit anti-AFP serum in the presence of 1 μg of carrier AFP (Mano & Chou, 1982) and analyzed by 10% polyacrylamide–SDS gel electrophoresis and fluorography.

Nucleic Acid Hybridization Analysis. Total RNA was isolated by the guanidinium thiocyanate/CsCl method (Chirgwin et al., 1979), and poly(A$^+$) RNA was obtained by oligo(dT)–cellulose chromatography (Aviv & Leder, 1972). Poly(A$^+$) RNAs were separated in 1.2% agarose gels containing 2.2 M electroblotting. RNAs were sized by using the RNA ladder obtained from Bethesda Research Laboratories.

Uniformly labeled, single-stranded DNA probes from either M13 or linearized plasmid vector were synthesized by primer extension in the presence of [α-^{32}P]dATP and dGTP (800 Ci/mmol; Amersham), purified by polyacrylamide–urea gel electrophoresis, and used in Northern blot hybridization. The filters were hybridized at 42 °C in a buffer containing 50% formamide, 10% dextran sulfate, and probe (10^6 cpm/mL) as previously described (Chou et al., 1988). RNA blots were washed 2 times in 300 mM NaCl/30 mM sodium citrate containing 0.5% SDS for 30 min each at room temperature and then 4 times in 15 mM NaCl/1.5 mM sodium citrate containing 0.1% SDS for 60 min each at 67 °C.

S1 Nuclease Analysis. The *Eco*RI/*Hin*dIII (nucleotides 1–1089 of rat AFP and 40 bp of the pUC18 polylinker) and the *Sph*I/*Pvu*II (nucleotides 680–1352) fragments of RAF5PP were used to examine the 5′ end and the middle of the AFP mRNA, respectively. Both fragments were end-labeled using T4 polynucleotide kinase, and the *Eco*RI/*Hin*dIII and *Sph*I/*Pvu*II probes were digested with *Pst*I and *Bam*HI, respectively, to yield 5′ end-labeled antisense strands. A uniformly labeled, single-stranded fragment (800 bases) generated by primer extension of RAF3PP was used to examine the 3′ end of the AFP mRNA. The 800-base fragment, released by *Nar*I digestion, contained 500 bases of rat AFP (nucleotides 1529 to the 3′ end) and 300 bases of M13mp18 sequence.

S1 nuclease mapping experiments were performed by hybridizing 40 μg of poly(A$^+$) RNA isolated from RLA209-15 cells with the respective probes (10^6 cpm) in buffer containing 80% formamide, 0.4 M NaCl, 0.04 M Pipes (pH 6.4), and 1 mM EDTA at 46 °C for 18 h. The hybrids were digested with S1 nuclease (900 units; Boehringer Mannheim Biochemicals) for 30 min at 42 °C and then electrophoresed on polyacrylamide–urea gels.

Primer Extension. The primer was a 30-base oligonucleotide (5′-GGGTAGTTTGCAGCACTCTGCTGTTTTGCT-3′) synthesized at the Laboratory of Molecular Pharmacology, Food and Drug Administration. The oligonucleotide was labeled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase and purified by electrophoresis on a 20% polyacrylamide gel. Primer extension reactions contained 40 μg of poly(A$^+$) RNA and 2 × 10^5 cpm of primer in hybridization buffer containing 10 mM Pipes, pH 6.4, 0.4 M NaCl, and 1 mM EDTA and were incubated at 65 °C for 3 h. The extension reaction was performed in the presence of 10 mM each of the dNTPs and 1000 units of AMV reverse transcriptase at 42 °C for 1 h. Samples were extracted with phenol/chloroform, ethanol precipitated, and analyzed on sequencing gels.

Isolation of a cDNA Clone Encoding the 1.7-kb AFP mRNA. An adult rat liver cDNA library in λgt11 was kindly provided by Dr. Frank Gonzalez (National Institutes of Health). A total of 4 × 10^6 plaques were screened with a nick-translated cDNA fragment, pRAF6 (Jagodzinski et al., 1981). The 1.2-kb *Eco*RI fragment of one positive clone (ARFP5) was subcloned into plasmid Gemini 1 to facilitate large-scale preparation. ARFP5ES was constructed by subcloning the *Eco*RI–*Sst*I fragment of ARFP5 (corresponding to nucleotides 1–186) into M13mp18.

DNA Sequence Analysis. ARFP5 and RAFGHH were sequenced by the Sanger dideoxy chain-termination method (Sanger et al., 1977) using deoxyadenosine 5′-(α-[^{35}S]thiotriphosphate) (400 Ci/mmol, Amersham).

RESULTS

1.7-kb AFP mRNA Encodes the Variant AFP. RLA209-15 fetal rat hepatocytes are temperature sensitive for maintenance of the differentiated fetal liver phenotype (Chou & Schlegel-Haueter, 1981). At either 40 or 33 °C, these cells produce both the fetal (69K and 73K) and variant (65K) AFPs and both the 2.2- and 1.7-kb AFP mRNAs (Chou et al., 1982; Chou & Savitz, 1986). However, RLA209-15 cells grown at 33 °C with a transformed phenotype produce mainly the 1.7-kb mRNA and the 65K AFP, suggesting that the 1.7-kb AFP mRNA encodes the variant AFP (Chou & Savitz, 1986). To verify this suggestion, we performed hybrid-selected translation using RNA isolated from RLA209-15 cells grown at 33 °C. RNA was hybrid selected with a single-stranded, antisense AFP cDNA, RAF3PP, which contains the fetal rat

AFP mRNA sequence extending from nucleotide 1529 to the 3′ end. Results in Figure 1A demonstrated that the hybrid-selected RNA directed the synthesis of two polypeptides of 67K and 50K which are immunoprecipitated by anti-AFP serum. The 67K and 50K polypeptides are the respective precursors of the fetal and variant AFPs (Chou & Savitz, 1986). Only trace quantities of anti-AFP-precipitable polypeptides were detected in the translation products directed by unbound RNA or RNA before hybrid selection (Figure 1A).

The anti-AFP serum used has been characterized extensively (Mano & Chou, 1982; Chou et al., 1982, 1988; Chou & Savitz, 1986). It was produced in rabbits by using a purified preparation of fetal rat AFP (Figure 1B) isolated from amniotic fluid of pregnant rats at 18-day gestation. It specifically precipitated the fetal AFP of 69K and 73K, and immunoprecipitation was completely prevented by the addition of excess fetal rat AFPs (Mano & Chou, 1982; Chou et al., 1982). This anti-AFP serum precipitated a 65K polypeptide in RLA209-15 cells, and the immunoprecipitation was also blocked by the addition of excess fetal rat AFPs (Chou et al., 1982, 1988; Chou & Savitz, 1986) (Figure 1C). Furthermore, the three precursor polypeptides (the 59K partially processed form, the 48K nonglycosylated form synthesized in the presence of tunicamycin, and the 50K in vitro translation product) of the 65K AFP and the 67K in vitro translation product of the fetal AFP were precipitated by this antiserum, and immunoprecipitation was prevented by excess fetal rat AFPs (Chou et al., 1982, 1988; Chou & Savitz, 1986) (Figure 1D,E). However, prebleeds from these rabbits did not precipitate any rat AFPs (data not shown). In addition to the antibody prepared in this laboratory, the 65K variant AFP and its 59K partially glycosylated precursor polypeptide could also be precipitated by two antibody preparations (bleeds of rabbit 11 and 12) obtained from Dr. J.-F. Chiu (University of Vermont). Data for antiserum of rabbit 11 are shown in Figure 1F.

Comparison of the 1.7- and 2.2-kb AFP mRNAs. In contrast to cells grown at 33 °C, at 40 °C RLA209-15 cells are phenotypically normal (Chou & Schlegel-Haueter, 1981; Chou et al., 1982) and produce mainly the 2.2-kb AFP mRNA (Figure 2). In order to differentiate between the 2.2- and 1.7-kb AFP mRNAs, we used hybridization probes specifying either the 5′, the middle, or the 3′ sequence of the fetal AFP mRNA. RAFGHA (containing the first exon Z) and RAF5PB (containing nucleotides 1–720) hybridized only with the 2.2-kb transcript (Figure 2). RAF5PS (containing nucleotides 1–954) hybridized strongly with the 2.2-kb RNA but weakly with the 1.7-kb RNA. These data indicate that the 1.7-kb AFP mRNA lacks the 5′ sequences of the 2.2-kb AFP mRNA. RAFPP (containing nucleotides 925–1529) and RAF3PP (containing nucleotide 1529 to the 3′ end), however, hybridized strongly with both the 2.2- and 1.7-kb RNAs, indicating that the 1.7-kb RNA contains sequences similar to the middle and the 3′ regions of the 2.2-kb AFP mRNA.

To demonstrate the sequence homology between the two RNAs, a series of S1 nuclease mapping experiments were performed. An 800-base uniformly labeled, single-stranded DNA fragment containing nucleotide 1529 to the 3′ end of the fetal AFP mRNA (500 bases, RAF3PP) and 300 bases of M13 vector sequence was used to map the 3′ regions of the two transcripts (Figure 3C). RNA from RLA209-15 cells grown at either 40 or 33 °C yielded only one 500-base protected fragment. This and the Northern blot results shown above indicate that the 2.2- and 1.7-kb AFP mRNAs are similar in the 3′ region.

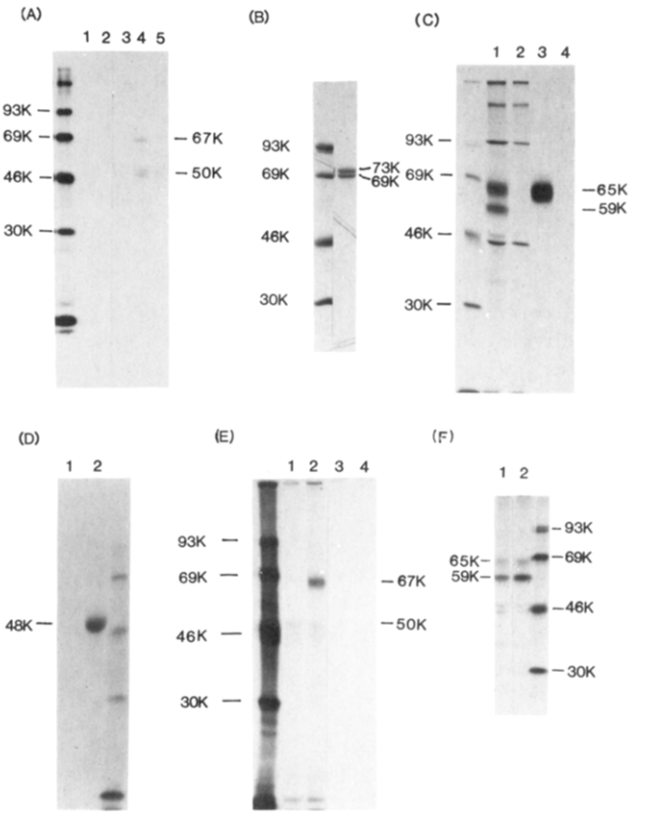

FIGURE 1: Cell-free synthesis of AFP directed by RNA selected by hybridization with an AFP cDNA and characterization of the rabbit anti-AFP antiserum. (A) Hybrid selection. One milligram of total RNA isolated from RLA209-15 cells grown at 33 °C was hybridized to the single-stranded RAF3PP in M13mp18 (antisense) immobilized on nitrocellulose filters. The hybrid-selected RNA was used to direct cell-free protein synthesis in a rabbit reticulocyte lysate system. AFPs in the in vitro translation products were isolated by immunoprecipitation and analyzed by polyacrylamide–SDS gel electrophoresis and fluorography. Lane 1, no RNA; lane 2, calf liver tRNA (2.5 μg); lanes 3–5, RNA from RLA209-15 cells: unbound RNA (2.5 μg, lane 3), hybrid-selected RNA (lane 4), and total RNA (2.5 μg, lane 5). (B) Polyacrylamide–SDS gel analysis of purified rat AFP used for immunization of rabbits. AFPs were stained with Coomassie brilliant blue. (C, D, and E) Competition for immunoprecipitation with excess amounts of purified rat AFPs. (C) RLA209-15 cells were grown at 33 °C and were labeled with L-[^{35}S]methionine for 3 h. The newly synthesized AFPs in cell lysates (lanes 1 and 2) and media (lanes 3 and 4), isolated by immunoprecipitation in the presence of either 1 μg (lanes 1 and 3) or 20 μg (lanes 2 and 4) of purified fetal AFPs, were analyzed by polyacrylamide–SDS gel electrophoresis and fluorography. (D) RLA209-15 cells were grown at 33 °C and were labeled with L-[^{35}S]methionine for 3 h in the presence of tunicamycin (1 μg/mL). Cells were pretreated with tunicamycin for 3 h prior to labeling. The newly synthesized AFPs in media were isolated by immunopreciitation in the presence of either 1 μg (lane 2) or 20 μg (lane 1) of purified fetal AFPs and analyzed by polyacrylamide–SDS gel electrophoresis. (E) Two sets of RLA209-15 cultures were grown at 33 °C initially. The first set of cultures was maintained at 33 °C throughout growth. The second set of cultures was shifted to 40 °C after 4 days of growth (day 0) at 33 °C. After an additional 3 days of growth at either 33 °C (lanes 1 and 3) or 40 °C (lanes 2 and 4), cells were lysed for RNA preparation. Three micrograms of poly(A^+) was used to direct cell-free protein synthesis in a rabbit reticulocyte lysate system. AFPs in the in vitro translation products were isolated by immunoprecipitation in the presence of either 1 μg (lanes 1 and 2) or 20 μg (lanes 3 and 4) of purified fetal AFPs and analyzed by polyacrylamide–SDS gel electrophoresis and fluorography. (F) Immunoprecipitation of newly synthesized variant AFP in RLA209-15 cells by an antiserum preparation of J.-F. Chiu. RLA209-15 cells were grown at 33 °C and were labeled with L-[^{35}S]methionine for 3 h. The newly synthesized AFPs in cell lysates were isolated by immunoprecipitation in the presence of 1 μg of purified fetal AFP using either our anti-AFP serum (lane 1) or an anti-AFP serum obtained from Dr. J.-F. Chiu (lane 2).

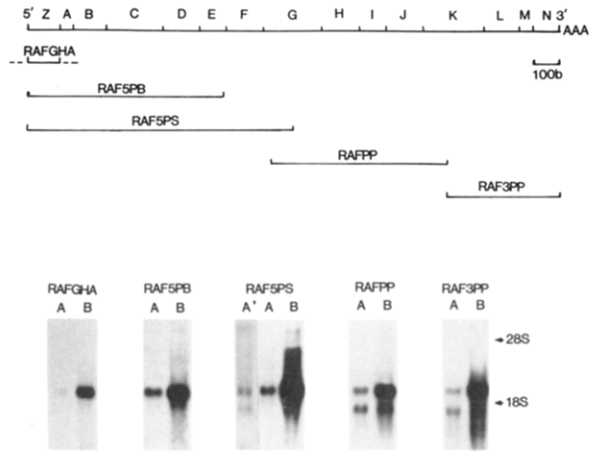

FIGURE 2: Analysis of AFP mRNAs in RLA209-15 cells grown at 33 and 40 °C. Two sets of cultures were grown at 33 °C initially. The first set of cultures (A or A′) was maintained at 33 °C (transformed phenotype) throughout growth. The second set of cultures (B) was shifted to 40 °C (differentiated phenotype) after 4-days growth (day 0) at 33 °C. After an additional 3-days growth at either 33 or 40 °C, cells were lysed for RNA preparation. Five micrograms of poly(A^+) RNA was applied to each lane and was separated by agarose gel electrophoresis. RNA was hybridized to ^{32}P-labeled AFP genomic or cDNA fragments RAFGHA, RAF5PB, RAF5PS, RAFPP, and RAF3PP as described under Materials and Methods. Z, A, B, C, D, E, F, G, H, I, J, K, L, M, and N are the 15 exons of the 2.2-kb AFP mRNA in their relative size (Jagodzinski et al., 1981; Turcotte et al., 1985).

A 5′ end-labeled *Eco*RI/*Hin*dIII probe was used to examine the 5′ end of the two AFP mRNAs. This probe was digested with *Pst*I, and the resulting *Hin*dIII site-labeled fragment was hybridized with RNA isolated from RLA209-15 cells, and the hybrid was analyzed by S1 nuclease digestion (Figure 3A). RNA isolated from RLA209-15 cells grown at 33 °C, which produced predominantly the 1.7-kb RNA, yielded mainly the 217-base DNA band by S1 nuclease analysis. RNA isolated from RLA209-15 cells grown at 40 °C yielded a major band of 1089 bases and several minor bands including the 217-base band. As expected, the 1089-base fragment mapped to the 2.2-kb AFP mRNA. The 217-base fragment, which was found in cells grown at either temperature, probably maps the 1.7-kb variant AFP mRNA. This analysis locates a region of divergence between the two AFP mRNAs at nucleotide 873 of the 2.2-kb RNA, which is the 5′ boundary of its eighth (G) exon.

To confirm these results, a 5′ end-labeled *Sph*I/*Pvu*II probe was digested with *Bam*HI, and the resulting 5′ *Pvu*II site-labeled fragment (containing nucleotides 725–1352 of the fetal AFP mRNA) was hybridized with RNA from RLA209-15 cells and analyzed by S1 nuclease digestion (Figure 3B). Two protected bands of 628 and 480 bases were obtained in cells grown either at 33 or 40 °C, although additional bands were found in cells grown at 40 °C. The 628-base band, which maps to the 2.2-kb AFP transcript, was present in greater amounts in cells grown at 40 °C than in cells grown at 33 °C. The 480-base band maps the 1.7-kb AFP mRNA to nucleotide 873 which is the beginning of the eighth exon of the 2.2-kb RNA.

Characterization of the 1.7-kb AFP mRNA by Primer Extension. Nuclease S1 analysis demonstrated that the 1.7-kb AFP mRNA contains similar sequences to the fetal AFP mRNA extending from the beginning of the eighth exon (nucleotide 873) to the 3′ end (about 1100 bases). However, the size of the variant AFP mRNA, estimated by Northern hybridization using an RNA ladder, was approximately 1700 bases, suggesting additional sequences at the 5′ end. To test this possibility, we performed primer extension experiments using a 5′ end-labeled 30-base oligonucleotide probe (corresponding to nucleotides 918–947 of the fetal AFP mRNA) and an RNA preparation containing predominantly the 1.7-kb AFP mRNA which was isolated from RLA209-15 cells grown at 33 °C (Figure 4B). Primer extension yielded two fragments of 480 and 225 bases at about equal intensity (Figure 4A). These fragments mapped to positions 405 and 150 bases 5′ to the beginning of the eighth exon (nucleotide 873) of the fetal AFP mRNA, suggesting that RLA209-15 cells contain two variant AFP mRNA species of approximately 1250 and 1500 bases in length. Since only one AFP-like RNA of 1.7 kb was detected and it is known that reverse transcriptase can produce incomplete copies which are often nonrandom in length (Efstratiadis et al., 1975; Payvar & Schimke, 1979), our results suggest that the size of this variant AFP mRNA without the poly(A) tail is about 1500 bases.

Isolation of a cDNA Encoding the Variant AFP mRNA. Although the 1.7-kb AFP mRNA was first demonstrated in the RLA209-15 fetal rat hepatocyte line, we found that this RNA is the major AFP mRNA species of adult rat liver (Chou et al., 1988). Thus, an adult rat liver cDNA library was screened with a nick-translated cDNA insert from pRAF6 which contains sequences shared by both the 2.2- and 1.7-kb AFP mRNAs. A total of 4×10^6 plaques screened yielded 12 positives which turned out to be identical after further analysis. The restriction map of one positive clone, ARFP5, and its 186 bp 5′ and 190 bp 3′ sequences are shown in Figure 5. The restriction map of clone ARFP5 (Figure 5A) was similar to that of the middle and 3′ regions of the cDNA encoding the fetal AFP mRNA (Jagodzinski et al., 1981). The 3′ sequences of ARFP5 and the 2.2-kb AFP mRNA are identical (Figure 5C). The 5′ sequence similarity between ARFP5 and the 2.2-kb RNA begins at nucleotide 873 of the 2.2-kb AFP mRNA which is the 5′ boundary of its eighth (G) exon (Figure 5B). ARFP5 contains an additional 90 bp sequence 5′ to the boundary of the G exon (the additional 9 bp 5′ sequence belongs to the *Eco*RI linker used for constructing the library), and this sequence is not present in the 2.2-kb RNA. The 87 bp 5′ sequence in ARFP5, which maps inside the G exon, contains a single base pair substitution. Nucleotide 893 of the fetal AFP mRNA is G, but the corresponding nucleotide in ARFP5 (nucleotide 120) is A. The deduced amino acid sequences of ARFP5 in three possible reading frames are shown in Figure 5B.

A uniformly labeled, single-stranded probe synthesized by primer extension of ARFP5ES (a subclone containing the 186 bp 5′ sequence of ARFP5) using a 20-base primer corresponding to nucleotides 74–93 shown in Figure 5B was used to map the 90 bp 5′ sequence of ARFP5 in the rat AFP gene (Figure 6). This probe hybridized only with genomic subclones carrying sequences of the seventh intron of the rat AFP gene, and this was confirmed by sequence analysis of clone RAFGHH (Figure 6B).

Probe containing the 84-base 5′ sequence of ARFP5 (nucleotides 10–93 shown in Figure 5B) hybridized only with the 1.7-kb AFP mRNA but not with the 2.2-kb fetal AFP mRNA (Figure 7B). However, a probe (the 186-base insert of ARFP5ES) containing an additional 87-base sequence of the G exon of the 2.2-kb AFP mRNA hybridized with both the 2.2- and 1.7-kb AFP mRNAs (Figure 7A).

A 1.7-kb AFP mRNA was also expressed in the CRL1548 rat hepatoma cell line (Figure 7). Like the 1.7-kb RNA of RLA209-15 cells, this hepatoma AFP mRNA hybridized to a probe containing the 1.7-kb AFP mRNA-specific sequence of ARFP5 (Figure 7B). Furthermore, the 1.7-kb RNA was

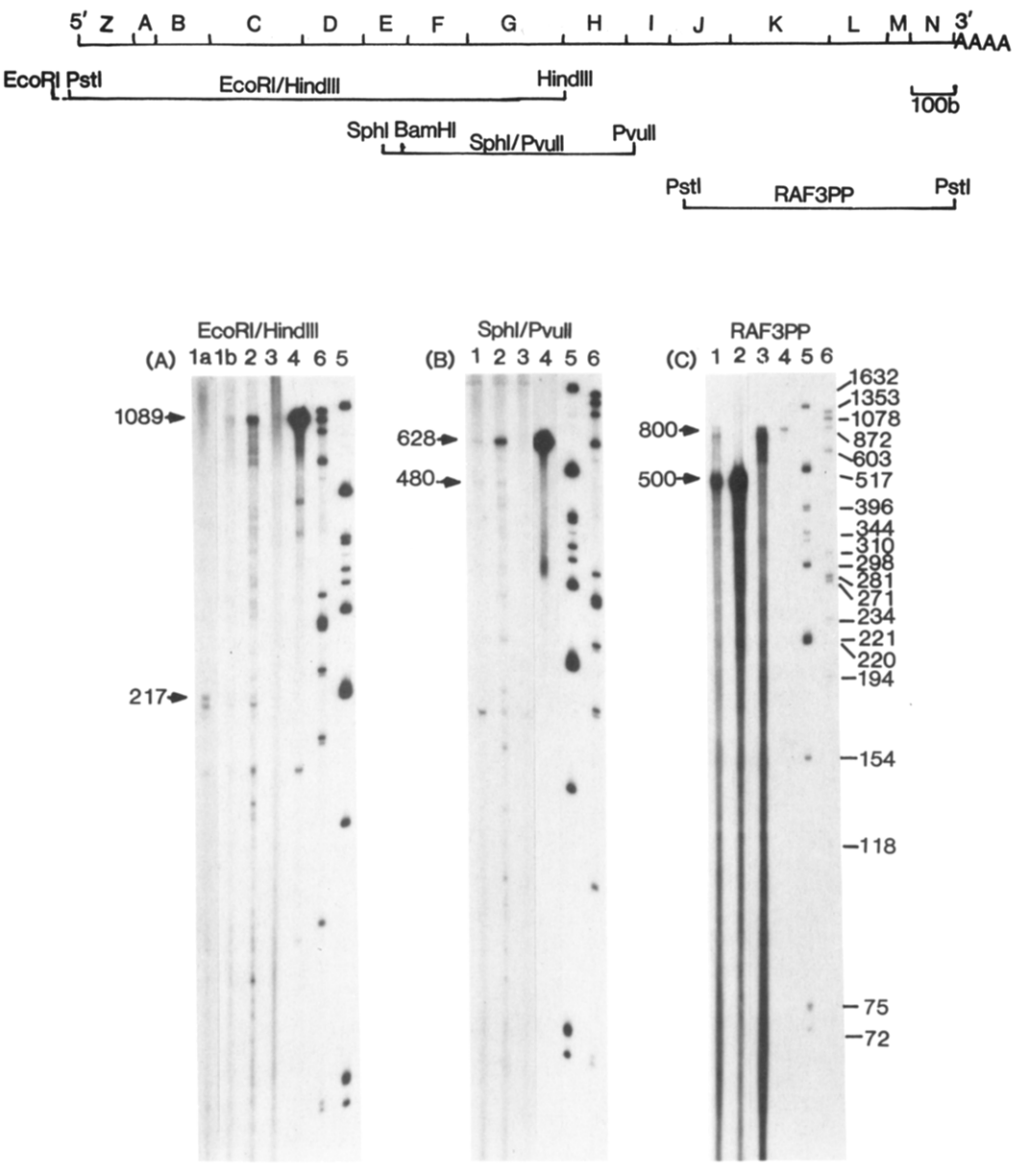

FIGURE 3: S1 nuclease mapping of AFP mRNAs in RLA209-15 cells grown at 33 and 40 °C. The culture conditions were as described in the legend to Figure 2. RNA was annealed to a 5′ end-labeled *Pst*I/*Hin*dIII fragment, a 5′ end-labeled *Bam*HI/*Pvu*II fragment, or a uniformly labeled RAF3PP fragment. These hybrids were digested with S1 nuclease as described under Materials and Methods. The S1 nuclease protected fragments were separated by electrophoresis on polyacrylamide–urea gels. Lane 1 (a and b are repeated experiments), poly(A^+) RNA (40 μg) from RLA209-15 cells grown at 33 °C; lane 2, poly(A^+) RNA (40 μg) from RLA209-15 cells grown at 40 °C; lane 3, calf liver tRNA (40 μg); lane 4, probe alone; lanes 5 and 6, ^{32}P-labeled markers: *Hin*fI digest of PBR322 (lane 5) and *Hae*III digest of $\phi X174$ (lane 6).

the only AFP mRNA detected in CRL1548 cells with either a probe containing the 186-base insert of ARFP5ES (Figure 7A) or a probe containing the middle (RAFPP) or 3′ sequence (RAP3PP) of the 2.2-kb AFP mRNA (data not shown). Our results indicate that the 1.7-kb AFP mRNAs expressed both in RLA209-15 and in CRL1548 cells are indistinguishable and CRL1548 cells do not express the 2.2-kb AFP mRNA.

Discussion

In the present study, we have characterized a 1.7-kb variant AFP mRNA produced by the RLA209-15 fetal rat hepatocyte line. This variant AFP mRNA possesses several characteristics. First, it encodes a 65K variant AFP. Second, the variant AFP mRNA contains sequences similar to the fetal AFP mRNA in the region extending from the 5′ boundary of the eighth exon (G, corresponding to nucleotide 873) to the 3′ end of the fetal AFP mRNA with the exception of a single base substitution. Third, the 1.7-kb AFP mRNA differs from the 2.2-kb fetal AFP mRNA in sequences at the 5′ end. The 1.7-kb RNA lacks sequences present in the first seven 5′ exons of the fetal AFP mRNA. Fourth, the 1.7-kb RNA contains an additional 5′ sequence which is not found in the fetal AFP mRNA.

Although normal adult rat liver produces only low levels of the 2.2-kb AFP mRNA, additional AFP mRNAs of smaller sizes have been observed by several laboratories (Gal et al., 1983; Sellem et al., 1984; Chou et al., 1988). We found that adult rat liver produces three AFP mRNAs of 2.2, 1.7, and 1.5 kb which are poly(A^+) (Chou et al., 1988). The 1.7-kb AFP mRNA in adult rat livers is indistinguishable from the variant AFP mRNA found in RLA209-15 cells (Y.-J. Y. Wan and J. Y. Chou, unpublished results). However, Petropoulos et al. (1985) have detected a 1.7-kb AFP mRNA mainly in the poly(A^-) fraction of adult rat liver RNA; this 1.7-kb RNA hybridized only with a probe containing the 5′ sequence but not with probes containing the middle or 3′ sequence of the fetal AFP mRNA. The experimental difference between these two laboratories is that Petropoulos et al. have isolated RNA from liver preparations that have been fractionated according to cell type and we have isolated our RNA from unfractionated

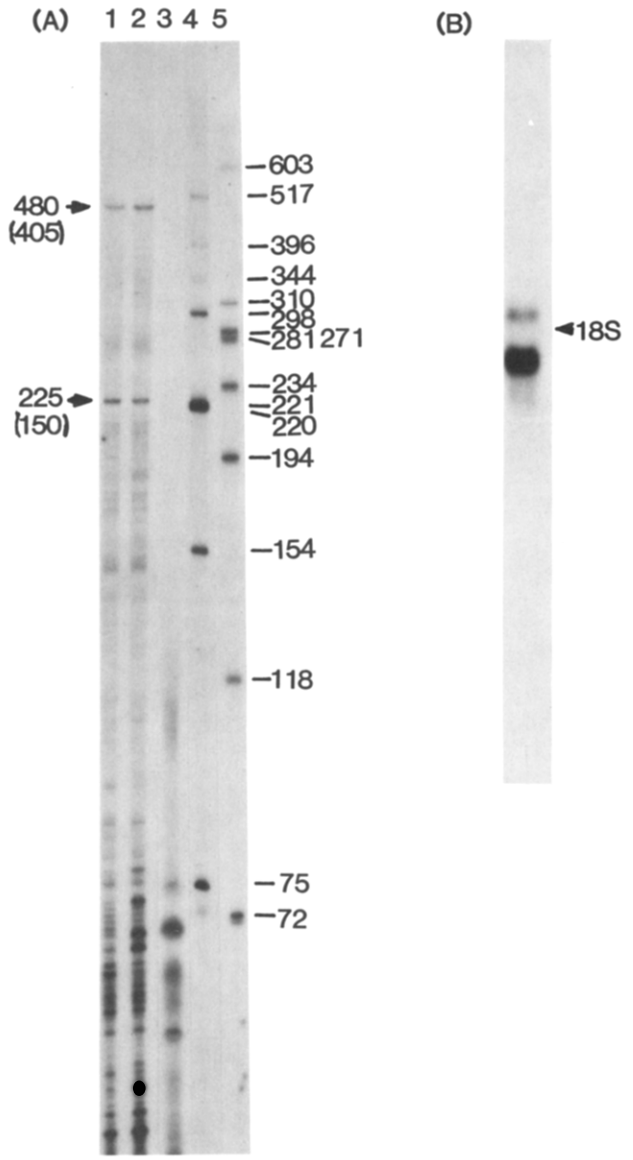

FIGURE 4: Mapping of the 5′ terminus of the 1.7-kb AFP mRNA by primer extension. RNA was isolated from RLA209-15 cells grown at 33 °C (transformed phenotype) which contained predominantly the 1.7-kb AFP mRNA (B). RNA was hybridized to the 5′ end-labeled 30-base primer (corresponding to nucleotides 918–947 of the fetal AFP mRNA) as described under Materials and Methods. The extended fragments were sized on a sequencing gel (A). Lanes 1 and 2, poly(A^+) RNA (40 μg) from RLA209-15 cells; lane 3, calf liver tRNA (40 μg); lanes 4 and 5, ^{32}P-labeled markers: *Hin*fI digest of PBR322 (lane 4) and *Hae*III digest of *φX174* (lane 5). Numbers in parentheses are lengths (bases) extended from the beginning of the eighth (G) exon of the fetal AFP mRNA.

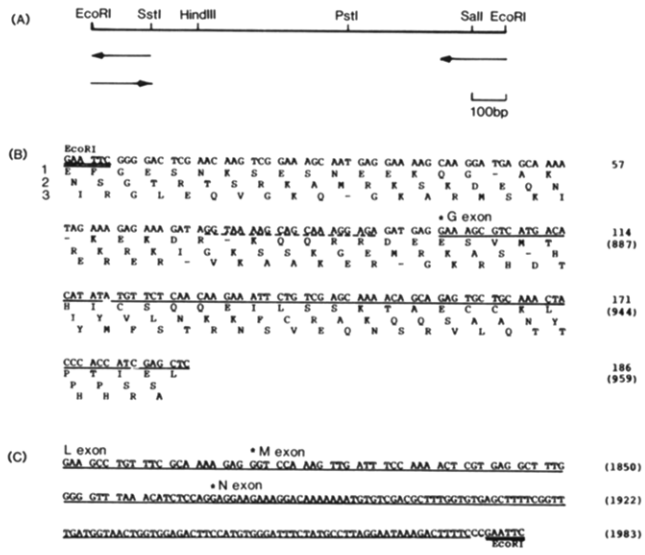

FIGURE 5: Restriction map and partial sequence analysis of ARFP5. Restriction map of ARFP5. Arrows indicate the direction and extent of nucleotide sequences determined in ARFP5. (B) The 5′-nucleotide sequences of ARFP5. The amino acid sequences in three possible reading frames are shown. Nucleotide 120 of ARFP5 is A, and the corresponding nucleotide in the fetal AFP mRNA is G (nucleotide 893). (C) The 3′-nucleotide sequences of ARFP5. Nucleotide sequences homologous to the 2.2-kb AFP mRNA are underlined, and G, L, M, and N are the corresponding exons. Numbers in parentheses are the corresponding nucleotide sequences of the 2.2-kb AFP mRNA.

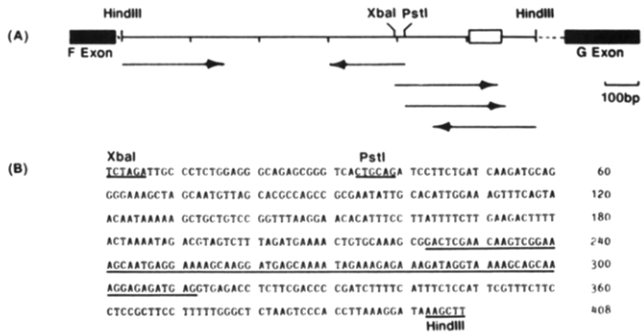

FIGURE 6: Restriction map and partial sequence analysis of rat AFP genomic clone RAFGHH. (A) Restriction map of RAFGHH. Arrows indicate the direction and extent of nucleotide sequences determined in RAFGHH. F and G are the seventh and eighth exons of the rat AFP gene in their relative size. The open box is the location of the 90 bp 5′ sequence of ARFP5 which is absent from the fetal AFP mRNA. (B) The 3′-nucleotide sequences of RAFGHH. The 90 bp 5′ sequence of ARFP5 is underlined.

whole livers. In addition, we could not detect a 1.7-kb AFP mRNA in the poly(A^-) fraction of adult liver RNA. At the present time, it is not known whether these could account for the different results obtained.

The similarity between the 1.7-kb AFP mRNAs of adult liver and RLA209-15 cells prompted us to screen an adult liver cDNA library for cDNAs encoding the variant AFP, and we report now the isolation and characterization of one such cDNA clone, ARFP5. Restriction endonuclease mapping and partial DNA sequencing analysis indicate that ARFP5 has extensive sequence similarity to the fetal AFP mRNA extending from nucleotide 873 (the beginning of the eighth exon G) to the 3′ end. However, this cDNA contains a 90 bp sequence 5′ to the G exon which is variant AFP mRNA specific and is located in the seventh intron of the rat AFP gene. Although the 5′ boundary of the variant AFP mRNA-specific sequence is unknown, the presence of sequence GT adjacent to the 3′ end of this sequence suggests that it is an alternative exon of the AFP gene.

The deduced amino acid sequences of the 5′ 177 bp nucleotide sequence determined in ARFP5 in three reading frames were examined. Both reading frames 1 and 3 contain three translation termination codons 5′ to the G exon. If reading frame 1 which is collinear with fetal AFP mRNA was the correct reading frame, then the 90 bp in ARFP5 is in the 5′ noncoding exon, and translation of the variant AFP starts within the G exon. It appears unlikely because the size of the variant AFP mRNA estimated both by Northern hybridization (1.7 kb) and primer extension (1.5 kb) is longer than ARFP5 (1.2 kb). In addition, the apparent molecular weight of the nonglycosylated form of the variant AFP is 48K, but the calculated molecular weight of the 328 amino acids encoded by ARFP5 is 36.8K.

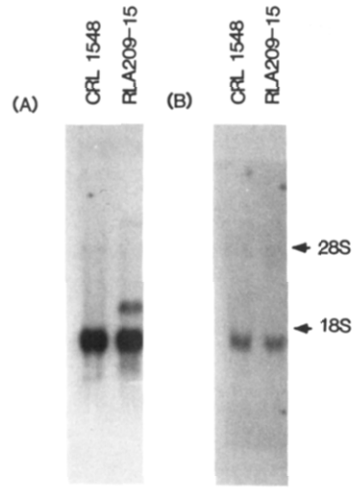

FIGURE 7: Hybridization of RNA in RLA209-15 and CRL1548 cells to ARFP5ES. Poly(A^+) RNAs were isolated from RLA209-15 cells grown at 33 °C (transformed phenotype) and CRL1548 cells grown at 37 °C. (A) RNAs were hybridized to a uniformly labeled, single-stranded probe containing the 186-base insert of ARFP5ES. This insert contains the 90-base 5′ sequence that is specific for the 1.7-kb AFP mRNA and the 87-base sequence present in the eighth exon of the 2.2-kb AFP mRNA. (B) RNAs were hybridized with a uniformly labeled, single-stranded probe prepared by primer extension of ARFP5ES with a 20-base primer (corresponding to nucleotides 74–93 shown in Figure 5B). This probe contains the 5′ sequences in ARFP5 which is specific to the 1.7-kb AFP mRNA.

The second reading frame in ARFP5 contains an open reading frame of 36 amino acid residues, but translation terminates within the G exon (nucleotide 110 in ARFP5, which corresponds to nucleotide 883 of the fetal AFP mRNA). The continued synthesis of the variant AFP from this reading frame would require translational frameshifting. Regulation of gene expression by frameshifting has been demonstrated in both prokaryotic organisms (Atkin et al., 1983; Dunn & Studier, 1983; Craigen et al., 1985; Craigen & Caskey, 1986) and eukaryotic systems (Jacks & Varmus, 1985; Clare & Farabaugh, 1985; Mellor et al., 1985; Wilson et al., 1986). However, it appears unlikely because frameshifting does not occur frequently and has been used only to control the expression of certain proteins that are required in small amounts (Bruce et al., 1986). In view of the preliminary nature of our cloning results, it is possible that a mistake in DNA sequencing results in the out of frame sequence in ARFP5.

We have unequivocally demonstrated the existence of two AFP mRNAs both in a fetal rat liver cell line and in adult liver. Furthermore, the 1.7-kb AFP mRNA is the only AFP transcript expressed in a rat hepatoma cell line. Since there is only one AFP gene per haploid genome in rat (Nahon et al., 1984), the fetal and variant AFP mRNAs may arise from the use of multiple sites of initiation, an alternative mode of exon splicing, or the combination of these events. Production of multiple transcripts from a single-copy gene by the above-mentioned mechanisms has been demonstrated in many other genes [for a review, see Leff and Rosenfeld (1986)]. It is known that the levels of these transcripts and the encoded proteins can vary considerably among different tissues (Amara et al., 1982; Rosenfeld et al., 1983; Schibler et al., 1983). We are currently examining the structure of the rat AFP gene and the mechanism responsible for generating the 1.7-kb variant AFP mRNA.

ACKNOWLEDGMENTS

We thank Drs. L. Belanger, L. L. Jagodzinski, T. D. Sargent, and J. Bonner for the gift of cDNA clones for rat AFP, Dr. F. Gonzalez for the gift of an adult rat liver cDNA library, Dr. J.-F. Chiu for the gift of anti-AFP antibodies, and Dr. J. Locker for the gift of genomic clone for rat AFP and his generosity in sharing with us the finding that CRL1548 rat hepatoma cells express a 1.7-kb AFP mRNA.

REFERENCES

Abelev, G. I. (1971) *Adv. Cancer Res. 14*, 295–358.
Abelev, G. I. (1974) *Transplant, Rev. 20*, 3–37.
Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S., & Evans, R. M. (1982) *Nature (London) 298*, 240–244.
Atkin, J. F., Nichols, B. P., & Thompson, S. (1983) *EMBO J. 2*, 1345–1350.
Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.
Bruce, A. G., Atkins, J. F., & Gesteland, R. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5062–5066.
Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J., & Rutter, W. J. (1979) *Biochemistry, 18*, 5294–5299.
Chou, J. Y., & Schlegel-Haueter, S. E. (1981) *J. Cell Biol. 89*, 216–222.
Chou, J. Y., & Savitz, A. J. (1986) *Biochem. Biophys. Res. Commun. 135*, 844–851.
Chou, J. Y., Ito, F., Evans, G., Chiu, J.-F., & Feldman, M. (1982) *Biochem. Biophys. Res. Commun. 108*, 1524–1530.
Chou, J. Y., Wan, Y.-J. Y., & Sakiyama, T. (1988) *Mol. Cell. Biol. 8*, 203–209.
Clare, J., & Farabaugh, P. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2829–2833.
Craigen, W. J., & Caskey, C. T. (1986) *Nature (London) 322*, 273–275.
Craigen, W. J., Cook, R. G., Tate, W. P., & Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3616–3620.
Dunn, J. J., & Studier, F. W. (1983) *J. Mol. Biol. 166*, 477–535.
Efstratiadis, A., Maniatis, T., Kafatos, F. C., Jeffrey, A., & Vournakis, J. N. (1975) *Cell (Cambridge, Mass.) 4*, 367–378.
Gal, A., Nahon, J., & Sala-Trepat, J. M. (1983) *Anal. Biochem. 132*, 190–194.
Gitlin, D., & Boesman, M. (1967) *J. Clin. Invest. 46*, 1010–1016.
Hirai, H. (1979) *Methods Cancer Res. 28*, 39–97.
Jacks, T., & Varmus, H. E. (1985) *Science (Washington, D.C.) 230*, 1237–1242.
Jacqueline, S., Miller, B. M., Paterson, R. P., Ricciard, L. C., & Bryan, E. R. (1983) *Methods Enzymol. 101*, 651–674.
Jagodzinski, L. L., Sargent, T. D., Yang, M., Glackin, C., & Bonner, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3521–3525.
Kafatos, F. C., Jones, C. W., & Efstratiadis, A. (1979) *Nucleic Acids Res. 7*, 1541–1552.
Kioussis, D., Eiferman, F., Van de Rijn, P., Gorin, M. B., Ingram, R. S., & Tilghman, S. M. (1981) *J. Biol. Chem. 256*, 1960–1967.
Leff, S. E., Rosenfeld, M. G., & Evans, R. M. (1986) *Annu. Rev. Biochem. 55*, 1091–1117.
Lehrach, H., Diamond, D., Wozney, J. M., & Boedtker, H. (1977) *Biochemistry 16*, 4743–4751.
Mano, T., & Chou, J. Y. (1982) *J. Biol. Chem. 257*, 5827–5830.
Meller, J., Fulton, S. M., Dibson, M. J., Wilson, W., Kingsman, S. M., & Kingsman, A. J. (1985) *Nature (London) 313*, 243–246.
Nahon, J.-L., & Sala-Trepat, J. M. (1984) in *Chromosomal Proteins and Gene Expression* (Reeck, G. R., Goodwin, G.

H., & Puigdomenech, P., Eds.) Vol. 101, pp 153-170, Plenum, New York.
Payvar, F., & Schimke, R. T. (1979) *J. Biol. Chem. 254*, 7636-7642.
Petropoulos, C. J., Yaswen, P., Panzica, M., & Fausto, N. (1985) *Cancer Res. 45*, 5762-5768.
Rosenfeld, M. G., Mermod, J. J., Amara, S. G., Swanson, L. W., Sawchenko, P. E., Rivier, J., Vale, W. W., & Evans, R. M. (1983) *Nature (London) 304*, 129-133.
Ruoslahti, E., Pihko, H., & Seppälä, M. (1974) *Transplant. Rev. 20*, 38-60.
Sakai, M., Morinaga, T., Urano, Y., Watanabe, K., Wegmann, T. G., & Tamaoki, T. (1985) *J. Biol. Chem. 260*, 5055-5060.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Schibler, U., Hagenbuchle, O., Wellauer, P. K., & Pittet, A. C. (1983) *Cell (Cambridge, Mass.) 33*, 501-508.
Sellem, C. H., Frain, M., Erdos, T., & Sala-Trepat, J. M. (1984) *Dev. Biol. 102*, 51-60.
Turcotte, B., Guertin, M., Chevrette, M., & Belanger, L. (1985) *Nucleic Acids Res. 13*, 2387-2398.
Wilson, W., Malim, M. H., Mellor, J., Kingsman, A. J., & Kingsman, S. M. (1986) *Nucleic Acids Res. 14*, 7001-7016.

cDNA Cloning of the *Octopus dofleini* Hemocyanin: Sequence of the Carboxyl-Terminal Domain[†,‡]

Walter H. Lang
Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331
Received February 22, 1988; Revised Manuscript Received May 17, 1988

ABSTRACT: A cDNA library was constructed in pUC 19, using poly(A+) RNA purified from *Octopus dofleini* branchial gland, which is the site of hemocyanin biosynthesis in cephalopods. The library was screened with an oligonucleotide probe derived from a portion of the partially known sequence of the C-terminal domain of *Paroctopus dofleini dofleini*. The clone with the longest insert—called pHC1—was sequenced and used as a probe for Northern blotting. It hybridized to a 9.5-kb RNA species, which was also visible as a band after ethidium bromide staining. The cDNA insert (approximately 1200 bp) of pHC1 contained an open reading frame of 1071 bp coding for 357 amino acids. In this insert, a region coding for 42 amino acids from the N-terminal end of the C-terminal domain is missing. These were obtained by sequencing a cloned primer extension product. By comparing our sequence with *Helix pomatia* β_c-hemocyanin unit D, we found 42.9% identical and 11.5% similar residues. One putative copper binding site (site B) was identified by homology to *Helix* hemocyanin and arthropodan hemocyanin. The location of a second possible site was identified.

Hemocyanins are copper-containing oxygen transport found in several orders of molluscs and arthropods (van Holde & Miller, 1982; Ellerton et al., 1983; Preaux & Gielens, 1984). Although molluscan and arthropodan hemocyanins are similar in amino acid composition, they differ greatly in molecular architecture: Arthropodan hemocyanins are composed of hexamers, which are found as (1 × 6)-, (2 × 6)-, (4 × 6)-, and (8 × 6)-meric aggregates, depending on the taxonomic group studied. There is one oxygen-binding site per subunit, which has a M_r of about 75000. On the other hand, molluscan hemocyanins are composed of 10 or 20 very large subunits (M_r about 350000–450000) that are arranged to form hollow cylinders. There are seven or eight oxygen-binding domains per subunit (M_r per domain of 45000–55000).

It has been of great interest whether these two hemocyanins evolved from a common ancestor or independently after the two phyla diverged. Their similarity in amino acid composition has been cited in favor of the first hypothesis, but this question can only be resolved by comparing amino acid sequences of both types. Currently, a number of amino acid sequences are known for subunits of arthropodan hemocyanins (Linzen et al., 1985; Nakashima et al., 1986; Bak & Beintema, 1987). Recently, the complete amino acid sequence for functional unit D of *Helix pomatia* β_c-hemocyanin was reported from amino acid sequencing of the isolated fragment (Drexel et al., 1987). This is the first sequence published for a functional domain of a molluscan hemocyanin.

When this sequence was compared to sequences of arthropodan hemocyanins, Drexel and co-workers (1987) found no similarities between the two types of hemocyanins except a region that corresponds to the "copper B" site in the hemocyanin of the arthropod *Panulirus interruptus*. This region is also found in known tyrosinase sequences. The N-terminal part of the *Helix* sequences contains sections that are clearly homologous to the tyrosinase sequences, but overall similarity is limited. The second copper-binding site could not be identified and must be completely different from the "copper A" site in arthropodan hemocyanins. The authors of this paper suggested on the basis of these findings that molluscan and arthropodan hemocyanins evolved independently from a common ancestral mononuclear copper protein. It is hypothesized that a gene duplication led to evolution of arthropodan hemocyanins, and a fusion with a different type of copper-binding structure led to the evolution of molluscan hemocyanins and tyrosinases.

[†] This work was supported by Grant DMB 85 10310 (to Kensal E. van Holde) from the National Science Foundation.
[‡] The nucleic acid sequence in this paper has been submited to GenBank under Accession Number J02835.

0006-2960/88/0427-7276$01.50/0