Pugsley, P. P., & Schwartz, M. (1985) *FEMS Microbiol. Rev. 32*, 3-38.
Righetti, P. G., & Chillenni, F. (1978) *J. Chromatogr. 157*, 243-251.
Smith, J. C., Derbyshire, R. B., Cook, E., Dunthorne, L., Viney, J., Biener, S. J., Sassenfeld, H. M., & Bell, L. D. (1984) *Gene 32*, 321-327.
Uhlén, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., & Lindberg, M. (1984) *J. Biol. Chem. 259*, 1695-1702.
Ullman, A. (1984) *Gene 29*, 27-31.
Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103-119.

Nucleotide Sequence of the Mouse α_1-Acid Glycoprotein Gene 1†

Robin Cooper,‡ D. Mark Eckley, and John Papaconstantinou*
Department of Human Biological Chemistry and Genetics, Division of Cell Biology, University of Texas Medical Branch, and Division of Cell Biology, Shriners' Burns Institute, Galveston, Texas Unit, Galveston, Texas 77550
Received January 28, 1987; Revised Manuscript Received April 27, 1987

ABSTRACT: In a previous paper we presented evidence for the existence of at least two α_1-acid glycoprotein (AGP) genes in the mouse. One of the cDNA clones characterized in those studies was used to isolate several unique AGP genomic clones. In these studies we present the complete sequence of one of the mouse AGP genes. The sequence analysis includes 595 base pairs (bp) 5′ to the site of initiation of transcription and 135 bp 3′ to the polyadenylation signal. This mouse AGP gene, designated AGP-1, has six exons, a structure similar to those of the AGP genes in rats and humans. Analysis of the sequence has revealed a number of potential regulatory sites. These include a run of alternating purine–pyrimidine bases [$(GT)_n$] at +2890 to +2945, flanked by three potential glucocorticoid receptor binding sites within intron 5. Two of these TGTTCT at +3069 to +3074 and +3082 to +3087 flank the $(GT)_n$ track at its 3′ end, and one, which is oriented in the opposite direction (AGAACA), at +2771 to +2776 flanks the track at its 5′ end. A longer version of the glucocorticoid receptor site, GGGTACAATGTGTCCT, has been located in the 5′ flanking region of the gene (−94 to −79); the sequence AGAACA is another potential glucocorticoid receptor site oriented in the opposite direction and located at −127 to −122. This entire region, from −146 to −42, in the mouse has a strong homology (~85%) to the 5′ flanking region of the rat AGP gene, which contains a 78-bp fragment (−120 to −42) that represents the minimal sequence required for glucocorticoid regulation. A sequence of 38 nucleotides (−22 to +16) that is homologous to similarly located sequences previously observed in three human acute-phase proteins has also been identified. We suggest that this sequence may represent an acute-phase protein regulatory element.

α_1-Acid glycoprotein (AGP), also known as orosomucoid, is a M_r 44 000 component of mammalian serum. In humans, the circulating amounts of AGP, and other acute-phase proteins, increase dramatically following a physiological insult such as acute inflammation, bacterial infection, major surgery, or burns (Koj, 1974). The biological significance of the acute-phase response is not well understood. Nevertheless, the reaction must have an important physiological role as the acute-phase proteins are highly conserved in evolution, and normal levels of these proteins are maintained even during severe malnutrition (Schmid, 1975; Ricca & Taylor, 1981; Ricca et al., 1982; Friedman, 1983).

The synthesis and secretion of the acute-phase proteins, including AGP, can be induced by injection of inflammatory agents such as turpentine. Both dexamethasone and a hepatocyte stimulating factor have also been shown to induce at least some of the acute-phase proteins (Ritchie et al., 1982; Baumann et al., 1983a,b, 1984a,b; Ritchie & Fuller, 1983). Studies with rats have indicated that, at the maximum point of induction by turpentine, the mRNA coding for AGP becomes one of the most abundant mRNA species in the liver (Northemann et al., 1983). Reports have been published in which both transcriptional and posttranscriptional mechanisms have been proposed for the regulation of the AGP gene expression in the rat (Vannice et al., 1984; Kulkarni et al., 1985). Changes in AGP mRNA pool levels, induced by dexamethasone in a rat hepatoma cell line, were reported to be due to stabilization of the primary AGP transcript and the efficient processing of this transcript to form a mature, functional cytoplasmic mRNA. This mechanism may be mediated by a protein factor whose synthesis is regulated by the glucocorticoid (Vannice et al., 1984). A more recent report on the effects of dexamethasone and turpentine in rats indicates that AGP regulation, at least in vivo, occurs at the transcriptional level (Kulkarni et al., 1985).

Previously, we reported the existence of at least two AGP genes in the mouse (Copper & Papaconstantinou, 1986). These conclusions were based on the sequence analysis of cDNA clones from a liver cDNA library. Our more recent work on sequence analysis of genomic clones shows the existence of an additional AGP gene in the mouse (unpublished observations). These data could explain the results of Baumann et al. (1984a,b), who observed multiple forms of AGP on two-dimensional protein gels. We wish, ultimately, to establish whether various mouse AGP genes are coordinately or differentially regulated in response to inducers such as inflammatory agents or hormones. As a first step toward this

† This research was supported by a grant from the Shriners' Burns Institute, Galveston, Texas Unit, to J.P.
* Correspondence should be addressed to this author at the Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch.
‡ Present address: Laboratory of Molecular Parasitology, The Rockefeller University, New York, NY 10021.

0006-2960/87/0426-5244$01.50/0

goal, we now present the complete sequence of one of the mouse AGP genes, i.e., mouse AGP-1. We also demonstrate the existence of a number of putative regulatory sites, including a run of alternating purine–pyrimidine bases [$(GT)_n$] flanked at its 3′ end by a pair of glycocorticoid receptor binding sites (TGTTCT) and at the 5′ end by a single receptor site that is oriented in the opposite direction (AGAACA). A longer version of this receptor binding site, GGGTACAATGTGTCCT, has been located in the 5′ flanking region of the gene. This site is flanked by another receptor binding site oriented in the opposite direction (AGAACA), which is located 27 nucleotides upstream. In addition, a highly conserved sequence of 38 nucleotides, localized between the TATAA box and the initiation codon, shows strong homology to the corresponding regions of other acute-phase protein genes (Dente et al., 1985).

Materials and Methods

Materials. Restriction enzymes were purchased from Bethesda Research Laboratories (Gaithersburg, MD), International Biotechnologies, Inc. (New Haven, CT), and Amersham Corp. (Arlington Heights, IL). The *Bal*31 exonuclease (fast form) was from International Biotechnologies, Inc. Calf intestinal phosphatase was from Boehringer Mannheim (Indianapolis, IN) or Pharmacia, Inc. (Piscataway, NJ). T4 DNA polymerase was from Pharmacia, Inc. T4 DNA ligase and *Escherichia coli* DNA polymerase I (Klenow fragment) were purchased from Amersham Corp. Dideoxy-DNA sequencing was carried out with sequencing kits from Bethesda Research Laboratories, Inc., or International Biotechnologies. Nick translation was carried out with a kit from Bethesda Research Laboratories. [^{32}P]dCTP (3000 Ci/mmol) and [α-^{35}S]deoxyadenosine 5′-(thiotriphosphate) was from Amersham Corp. or New England Nuclear (Boston, MA). Nitrocellulose was from Millipore Corp. (Bedford, MA) or Schleicher & Schuell (Keene, NH).

Isolation of Genomic Clones. A Balb/c mouse embryo library, generated by a partial *Mbo*I digestion and insertion into the *Bam*HI cloning site of Charon 28 (a gift from Dr. Philip Leder), was screened for AGP sequences with a mouse cDNA clone (pMAGP4) previously isolated (Cooper & Papaconstantinou, 1986). Approximately 1.1×10^6 phage were plated onto five large pans, transferred to nitrocellulose, alkali treated, and neutralized, and the DNA was fixed to the filter as described (Benton & Davis, 1977). The insert from an AGP cDNA clone (pMAGP4) was isolated and purified over a NACS column (Bethesda Research Laboratories) and labeled with ^{32}P by nick translation. This probe was hybridized to the filters in 50% deionized formamide, 5× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate), 50 mM phosphate buffer (pH 6.5), 200 μg/mL sheared and denatured salmon sperm DNA, Denhardt's solution [0.02% each of ficoll, poly(vinylpyrrolidone), and bovine serum albumin], and 1×10^8 cpm of probe for 36 h at 42 °C. The filters were washed 4 times at room temperature in 2× SSC–0.1% SDS and 3 times at 50 °C in 0.1× SSC–0.1% SDS with an abundant amount of each solution. The filters were then dried and autoradiographed for 2 days on Kodak X-omat film with a Du Pont Cronex Lightning Plus intensifying screen at −90 °C. Positive signals were rescreened at lower plaque densities until single plaques could be identified.

Analysis of DNA from Genomic Clones. DNA from the genomic clones was prepared according to the method of Enquist et al. (1979) except that 30 mM $MgSO_4$ was included in the L broth. Restriction enzyme digestions were carried out following the supplier's recommendations. In general, 1 μg of DNA was digested with 10 units of enzyme for 1–3 h at 37 °C. One-half volume of 25% glycerol, 5% SDS, and 0.5% bromophenol blue was added to stop the reactions. Gel electrophoresis was carried out in 1% agarose in TAE [40 mM tris(hydroxymethyl)aminomethane (Tris), 5 mM sodium acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.8] at 40 V for approximately 16 h. DNA was transferred to nitrocellulose as previously described (Cooper et al., 1984). Hybridizations were carried out as described above. *Bam*HI restriction fragments were subcloned into the plasmid pBR322, and DNA was prepared according to a procedure described by Norgard et al. (1979).

Generation of Bal31 Deletion Clones for Sequencing. Sequencing of the two *Bam*HI fragments containing one of the AGP genes was approached according to a technique to generate *Bal*31 deletion clones (Poncz et al., 1982) with modifications by Alsip and Konkel (1984). Briefly, the pBR322 subclones containing the fragments of interest were initially cleaved with a restriction enzyme that cut only once within the pBR322 portion of the plasmid and close to the site of insertion. The exonuclease *Bal*31 was then used to digest the linear DNA molecule for 1–10 min. Approximately 0.42 pmol of DNA per time point was digested with 0.19 unit of *Bal*31 (fast) in the buffer provided at 30 °C. These conditions yielded an approximate rate of digestion of 150 bp min^{-1} end^{-1}. The reactions were stopped by adding ethylene glycol bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA) up to 15 mM and carrier tRNA, purified over a Cellex D column, followed by a phenol–chloroform (1:1) extraction. At this stage, 5′ or 3′ protruding ends were filled in with T4 DNA polymerase by incubation at 37 °C for 10 min in 33 mM Tris–acetate (pH 7.9), 66 mM potassium acetate, 10 mM magnesium acetate, 0.1 mM dNTPs (dGTP, dATP, dTTP, dCTP), 0.5 mM dithiothreitol, 0.1 mg/mL bovine serum albumin (BSA), and 1.0 unit of T4 DNA polymerase. The reactions were stopped by the addition of EDTA to a final concentration of 20 mM. The processed DNAs were then digested with *Bam*HI and a second restriction enzyme that cut within the pBR322 portion of the plasmid at only one site. This combination of restriction digests allowed us to clone only the progressively shortened insert fragments (blunt and *Bam*HI cohensive ends) into a bacteriophage M13 vector (Yannisch-Perron et al., 1985) cut with *Bam*HI and *Hin*cII so that the processed (blunt) end lay proximal to the primer annealing site. A similar strategy was employed to sequence the opposite strand. Some additional M13 clones necessary for obtaining a complete sequence in both directions were generated by cloning particular restriction fragments.

DNA Sequencing. Sequencing was carried out by the Sanger chain termination method and with [α-^{35}S]thio-dATP as a label, a gradient gel being used (Sanger et al., 1977; Biggin et al., 1983; Alsip & Konkel, 1984). Sequence data were analyzed by a computer program designed by Dr. James M. Pustell, Harvard University.

Nuclease S1 Analysis. The 5′ transcription initiation site of the AGP-1 gene was determined by nuclease S1 mapping of AGP RNA. Briefly, a restriction fragment that included 633 bp 5′ to the ATG encoding the first methionine residue and terminating at a *Hin*cII site within the first exon was isolated. The 5′-phosphate groups were removed by treatment with calf intestinal alkaline phosphatase (Maniatis et al., 1982), and the fragment was then end-labeled with [^{32}P]dATP by use of T4 polynucleotide kinase (Amersham Corp.) (Maniatis et al., 1982). This labeled fragment was then annealed to partially purified poly(A)-containing RNA obtained from a

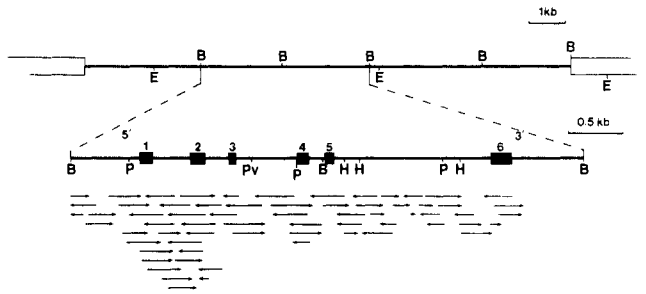

FIGURE 1: Partial restriction map of mouse AGP-1 gene (MAGP31). The solid black lines represent the insert into the *Bam*HI site of Charon 28. The short (left) and long (right) arms of Charon 28 are represented by the open boxes. Exons are represented by solid boxes. The complete sequencing strategy is shown, where arrows represent the extent and direction of each sequencing reaction. Restriction enzymes: B = *Bam*HI, E = *Eco*RI, P = *Pst*I, Pv = *Pvu*II, and H = *Hin*dIII.

turpentine-injected mouse under conditions that promoted the formation of DNA–RNA hybrids but minimized DNA–DNA hybrids (Maniatis et al., 1982; Favaloro et al., 1980) [80% deionized formamide, 0.4 M NaCl, 1 mM EDTA, 40 mM 1,4-piperazinediethanesulfonic acid (PIPES), pH 6.4, at 50 °C for 3 h after first heating to 80 °C for 15 min to denature the nucleic acids]. The annealed nucleic acids were then treated with nuclease S1 (400 units/mL) for 30 min at 37 °C (Maniatis et al., 1982). The reaction was stopped with EDTA, the solution extracted with phenol–chloroform (1:1), and the DNA precipitated with ethanol. The DNA pellet was dissolved in 10 μL of 5 mM Tris-HCl, pH 7.5, 10 mM EDTA, 50% deionized formamide, and 0.05% bromophenol blue, boiled for 3–5 min, and loaded onto a buffer gradient polyacrylamide sequencing gel and electrophoresed 2.5 h at 35 mA. The gel was dried and exposed with XAR film for 5–7 days at −70 °C with an intensifying screen.

RESULTS

Isolation and Sequencing of Mouse AGP-1 Gene. In a previous publication we provided evidence, based on sequence analysis of cDNA clones, for the existence of at least two AGP genes in the mouse (Cooper & Papaconstantinou, 1986). One of the cDNA clones (pMAGP4) was used to screen a mouse genomic library for AGP sequences. A total of five unique clones were identified, one of which, MAGP31, was initially chosen for further study. Southern blotting of various restriction digests and subclones (data not shown) gave rise to a restriction map for MAGP31 (Figure 1). These experiments suggested that the entire coding region of one AGP gene is within the 2.3- and 2.4-kb *Bam*HI fragments of MAGP31. Both of these fragments were sequenced with exonuclease *Bal*31 to generate an overlapping series of deletion fragments that could then be selectively cloned into appropriate M13 sequencing vectors (Poncz et al., 1982). In most regions more than two templates, from opposite DNA strands, were used to deduce the correct sequence. A diagram of the complete sequencing strategy can be seen in Figure 1. Our data establish the sequence for one of the mouse AGP genes, which we named AGP-1 (Figure 2). It includes 595 bp 5′ to the mRNA initiation site and 135 bp 3′ to the polyadenylation signal, AATAAA. The mouse AGP-1 gene has six exons and five introns, a structure exactly analogous to both rat (Liao et al., 1985; Reinke & Feigelson, 1985) and human (Dente et al., 1985) genes. Each of the introns follows consensus rules for splice junctions and lariat branch point sequences (Breathnach & Chambon, 1981; Padgett et al., 1984). The sequence of the exons agrees exactly with the sequence we reported for the cDNA clone pMAGP4 (Cooper & Papaconstantinou, 1986), with the exception of two differences at positions +1805 and +3237. In the cDNA we observed a T at each position, while in the genomic sequence we find a C at each position. Both these differences occur at the third position of the codon at each point. In both cases the amino acid encoded remains the same despite the single nucleotide transitions. This minor divergence in sequence could be attributed to strain differences in the mice used to generate each library. The cDNA library was derived from SWR/J mice while the genomic library was generated from Balb/c mice.

AGP-1 Transcriptional Initiation Sites. The site of initiation of mRNA synthesis was determined by nuclease S1 analysis. A restriction fragment was isolated that included the entire 5′ flanking region and that terminated at a *Hin*cII site within the first exon. Nuclease S1 analysis using this fragment annealed to partially purified liver poly(A)-containing RNA gave rise to fragments 109, 110, 115, 116, 117, and 118 nucleotides in size (Figure 3), placing the putative initiation sites at the positions underlined by equal signs beginning 38 bp 5′ to the ATG in Figure 2. The multiple bands observed could be explained by hypothesizing that there are two initiation sites or that the AGP mRNA is transcribed from two different genes with slightly different initiation sites. We cannot at this time distinguish between these two possibilities. Interestingly, in the rat AGP gene two transcriptional start sites, four bases apart, have been found (Reinke & Feigelson, 1985). Hence, two start sites in the mouse AGP-1 gene is a distinct possibility. We propose that the major site of mRNA initiation in this gene is that marked by the strongest, i.e., the 110-nucleotide, band. This nucleotide corresponds to an A and has been numbered +1 in Figure 2. This agrees with the observation that most eucaryotic mRNAs initiate transcription with an A residue (Breathnach & Chambon, 1981).

Putative Regulatory Sites in Mouse AGP-1 Gene. Examination of the sequence 5′ from the initiation site reveals a TATAA box at position −33 and a possible CAAT box, CCAAG, at position −103. Because AGP is induced by glucocorticoids, we searched the AGP-1 gene sequence for the glucocorticoid receptor binding hexamer TGTTCT or its reverse complement AGAACA (Slater et al., 1985; Karin et al., 1984) and found it present 7 times at positions −127, +101, +975, +2213, +2771, +3069, and +3082 (Figure 2). Three of these sequences, at position +2771 and a pair at +3069 and +3082, separated by only seven bases, flank a long GT-rich region in intron 5. Other acute-phase proteins have similar, though often shorter, alternating purine–pyrimidine-rich regions (Figure 4). These include the human haptoglobin related gene pair (Maeda, 1985), the human C reactive protein gene (Lei et al., 1985; Woo et al., 1985), the rat AGP gene (Reinke & Feigelson, 1985; Liao et al., 1985), and the human α_1-antitrypsin gene (Long et al., 1984).

In a recent study Baumann and Maquat (1986) identified a minimal sequence in the 5′ flanking region of the rat AGP gene extending from −42 to −120 bp that is required for glucocorticoid regulation. A comparison of this sequence with that of the mouse AGP-1 gene (−47 to −127 bp) indicates that they are highly conserved, i.e., showing an ~85% homology (Figure 5). We propose that this region of the mouse AGP-1 gene may be a putative glucocorticoid regulatory element. Our hypothesis is further supported by our sequence data, which reveal a reverse complement of the glucocorticoid receptor binding hexamer AGAACA at −122 to −127 bp. Furthermore, within this same region, beginning at site −94, the mouse AGP-1 sequence GGGgAgAATGTGcCag is 69% homologous to a longer version of the glucocorticoid receptor binding site

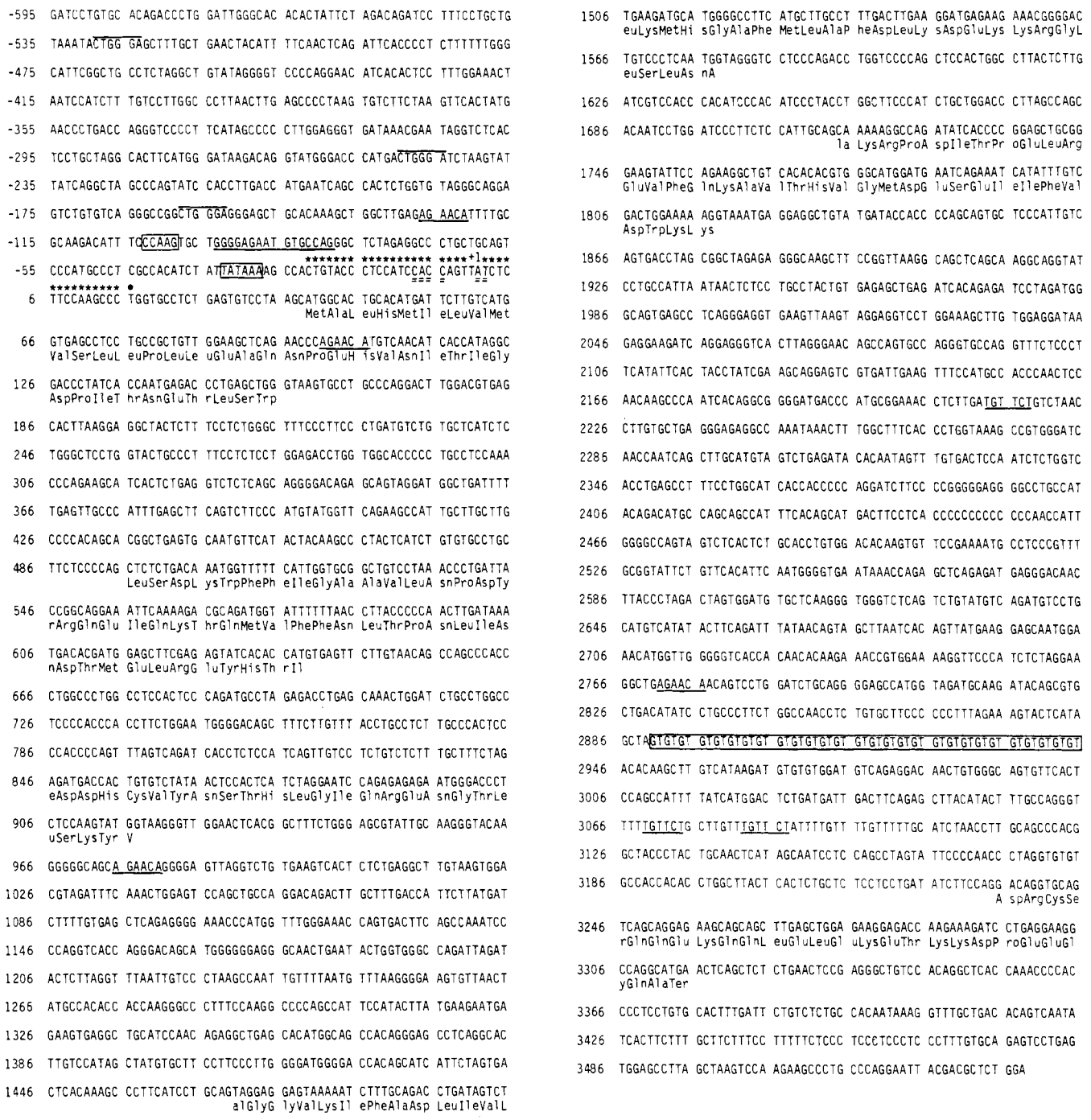

FIGURE 2: Complete nucleotide sequence of mouse AGP-1 gene derived from the insert in MAGP31. The nucleotide sequence of the AGP-1 gene including the 5′ and 3′ flanking sequences is shown. The amino acid sequence encoded by all six exons is shown below the corresponding nucleotide sequence. The sequence is numbered relative to the mRNA initiation site indicated by +1. Other possible initiation sites are underlined by equal signs. Potential glucocorticoid receptor binding sites are underlined. The TATAA box, putative CAAT box, and GT-rich region in intron 5 are enclosed in boxes. The region of homology to the consensus sequence described by Dente et al. (1985) is marked by asterisks above the sequence. The six-nucleotide core sequence found in the rat fibrinogen genes (Fowlkes et al., 1984) is marked by lines above the sequence.

whose consensus sequence is GGGTACAATGTGTCCT (Slater et al., 1985; Karin et al., 1984). This sequence has been located within the 5′ flanking region of several glucocorticoid sensitive genes, including mouse α-fetoprotein (Scott & Tilghman, 1983), mouse albumin (Izban and Papaconstantinou, unpublished results), rat AGP (Ricca & Taylor, 1981; Ritchie et al., 1982), and rat β-fibrinogen (Fowlkes et al., 1984). Whether these sequences actually function as glucocorticoid receptor sites in the mouse must yet be determined by direct assay.

Dente et al. (1985) have previously observed a conserved DNA sequence in the 5′ untranslated region of three human acute-phase proteins: haptoglobin, α_1-antitrypsin, and AGP. We have searched both mouse and rat AGP genes for this conserved DNA sequence. In each case a closely homologous sequence was found in analogous positions to that found in the human gene (Figure 6). Although the significance of this region is not known, the high degree of homology between the two rodent genes and the three human genes is noteworthy, especially because in all five genes the homology is located between the TATAA box and the methionine initiation codon and includes the transcriptional initiation site. We have not found homologous sequences in the mouse albumin or α-fetoprotein genes, both of which are non-acute-phase, liver-

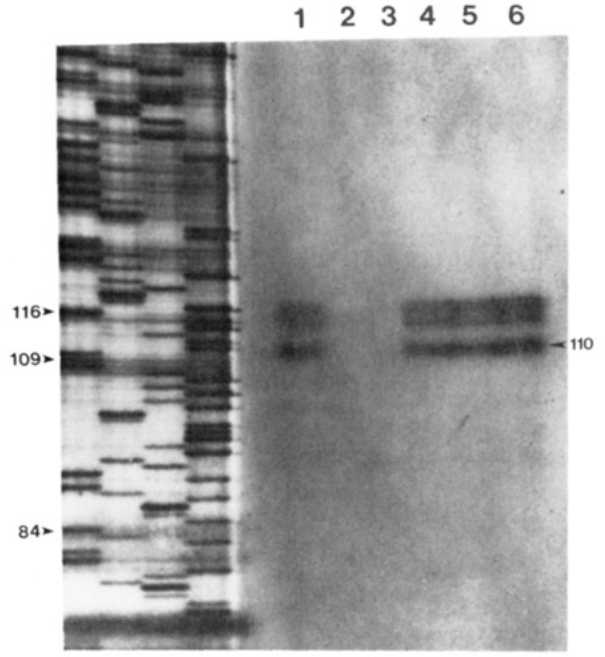

FIGURE 3: Nulclease S1 analysis of AGP mRNA. Different RNA samples were hybridized to a restriction fragment from the 5′ end of the AGP gene and treated with nuclease S1. The product was run alongside a sequencing reaction for molecular weight determination. Lanes 1, 4, 5, and 6 contain 5, 5, 10, and 20 μg, respectively, of liver poly(A)-containing RNA from a turpentine-injected mouse. Lane 2 contains 5 μg of liver poly(A)-containing RNA from an untreated mouse. Lane 3 contains no RNA other than the same amount of tRNA used as a carrier in the other reactions. Molecular weight markers are in nucleotides.

specific proteins. In the case of albumin, however, a negative acute-phase response has been reported in which transcription of this gene is repressed during an acute-phase reaction (Kulkarni et al., 1985). Thus, this sequence may be a positive regulatory element, specific for acute-phase reactants.

Mouse AGP Amino Acid Sequences. We have previously published a partial amino acid sequence for mouse AGP deduced from a partial cDNA clone (Cooper & Papaconstantinou, 1986). The sequence of the entire gene has allowed us to complete the amino acid sequence for one of the AGP proteins in the mouse and compare it to rat and human AGP (Figure 7). As the N-terminus of mature mouse AGP is not known, we have arbitrarily designated Gln as the N-terminal amino acid by analogy with the known mature rat amino acid sequence (Nagashima et al., 1981). The predicted N-terminus of the unprocessed AGP-1 protein sequence contains a hydrophobic region typical of secretory leader signal peptides. However, the exact length of the leader peptide is unclear as three in-phase initiation condons exist within the 5′ transcribed region. Homology to the rat sequence and the general rule that the 5′-most AUG codon begins translation suggest nevertheless an 18-residue signal peptide. Since AGP is a highly glycosylated protein, it was of interest to localize potential carbohydrate attachment sites in the mouse amino acid sequence as indicated by the presence of the sequence Asn-X-$^{Thr}_{Ser}$ (Eylar, 1966). Six potential sites can be found in the mouse sequence at positions 7, 16, 51, 58, 76, and 86. The same number are found in the rat sequence, all but one of which (at position 116) at identical positions. Human AGP is known to have five N-linked glycosylations, three of which are common to potential sites in the mouse and the rat (positions 16, 76, and 86). This conservation in number and position of carbohydrate side chains argues that they are probably important in the function of the protein.

Discussion

We have isolated a genomic clone that includes the entire sequence of one of the mouse AGP genes. This gene, which we named mouse AGP-1, has been sequenced from −595 bp 5′ to the site of initiation of transcription to 135 bp 3′ to the polyadenylation signal. Our sequence analysis has shown that the mouse gene has six exons and five introns and compares well with the sequence of a partial cDNA clone of mouse AGP (Cooper & Papaconstantinou, 1986) and the complete mRNA and gene sequence of rat AGP (Reinke & Feigelson, 1985; Liao et al., 1985). This gene structure has been well conserved, as is indicated by its homology and organizational similarity to the rat and human AGP genes (Dente et al., 1985).

We have previously provided evidence for the existence of at least two AGP genes in the mouse (Cooper & Papaconstantinou, 1986). The sequence of the gene reported here is in exact agreement with one of the cDNA clones we isolated previously, except for two single nucleotide differences attributable to minor variations in the two strains of mice used to generate the cDNA and genomic libraries.

Our analysis of the AGP-1 gene sequence has revealed a number of putative regulatory elements. These include several consensus glucocorticoid receptor binding sites, both in the 5′

Gene	Alternating Pu/Py rich region
Mouse AGP-1	3485 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540 ---- 3664 TGTTCTGCTTGTTTGTTCT 3682
Rat AGP	2652 GTGTGTGTGTGT 2663 --------- 2821 TGTTCTGTTTTGTTTGTTGT 2841
Human haptoglobin	2433 GTGTGTGTGTATGCAGTGTGTGTGTGTGTGTACATGCATGTGTGTGT 2482
Human haptoglobin related gene	14134 GTGTGTGTGTATGCATGTGTGTGTGTGCGTGTGTGTGTGTGTACATGCCTGTGTGTGT 14193
Human C-reactive protein	419 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGGTGTGT 455
Human alpha-1-antitrypsin	8296 GTGTATGTGTGTGCA 8310

FIGURE 4: Compilation of alternating purine–pyrimidine-rich regions from mouse AGP-1 gene and other acute-phase protein genes. Potential glucocorticoid receptor binding sequences have been underlined. References: rat AGP (Reinke & Feigelson, 1985; Liao et al., 1985), human haptoglobin (Lei et al., 1985; Woo, et al., 1985), human haptoglobin related gene (Maeda, 1985), and human C reactive protein [human α_1-antitrypsin (Long et al., 1984)].

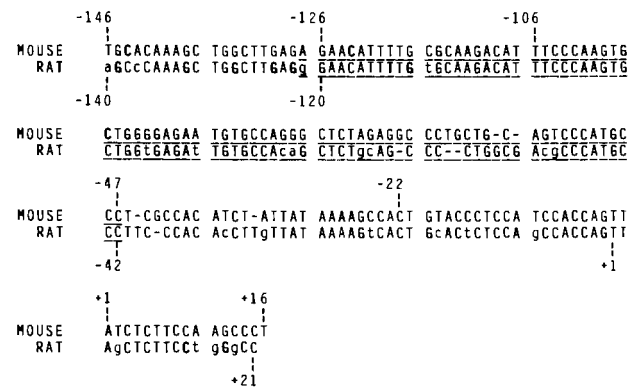

FIGURE 5: DNA sequence homology between the 5′ flanking region of the mouse (AGP-1) and the 5′ flanking region of the rat AGP gene that contains the glucocorticoid regulatory element. The underlined region from nucleotide −120 to nucleotide −42 in the rat sequence represents the minimal sequence required for glucocorticoid regulation reported by Baumann and Maquat (1986). The homologous sequence of the mouse AGP-1 gene ranges from −126 to −47. The capital letters depict regions of homology between the mouse and rat sequences, the lower case letters depict regions that lack homology, and the dash marks indicate regions of deletions or insertions. The site of initiation of transcription is depicted by +1.

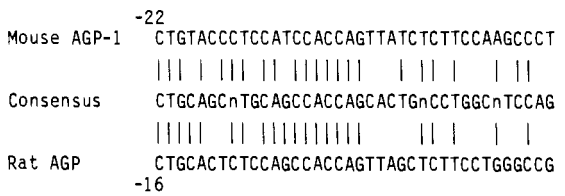

FIGURE 6: DNA sequence homology between the mouse and rat AGP genes and the consensus acute-phase sequence described by Dente et al. (1985). The numbers indicate the position with respect to the site of mRNA initiation. Vertical lines indicate a perfect match to the consensus sequence.

flanking region and within intron 5. The latter are found in close association with a long purine–pyrimidine-rich region composed of alternating GTs. It has been demonstrated that alternating purine–pyrimidine sequences can form Z-DNA, a left-handed helical form of DNA, which has been postulated to be a potential site for gene regulation (Rich et al., 1984) but whose natural occurrence and possible function in the regulation of gene expression are still controversial (Marx, 1985). Alternating purine–pyrimidine tracks have also been demonstrated to enhance the expression of genes cloned into recombinant plasmids (Hamada et al., 1984). Although $(GT)_n$ tracks exist in non-acute-phase genes (Nishioka & Leder, 1980; Shen et al., 1981; Miesfield et al., 1981; Hentschel, 1982) in telomeres (Walmsley et al., 1983) and as multiple copies in the human genome (Hamada & Kakuraga, 1982), it is possible that they may be nonspecific enhancers that can be modulated by adjacent specific regulatory sequences such as the glucocorticoid receptor sites.

Recently it has been shown that the region −42 to −120 bp of the rat AGP gene mediates maximal dexamethasone induction of the bacterial chloramphenicol acetyltransferase gene when transfected into mouse L-cells in transient expression assays (Baumann & Maquat, 1986). Our comparison of the sequence of the analogous region in the mouse AGP-1 gene clearly shows a strong homology between these regions, indicating that this may be a potential glucocorticoid regulatory element in the mouse AGP-1 gene. Analysis of the rat sequence, however, showed that this region does not contain sequence homologies to known glucocorticoid receptor binding sites, leading the authors to speculate that this sequence may represent a new type of glucocorticoid responsive element. In

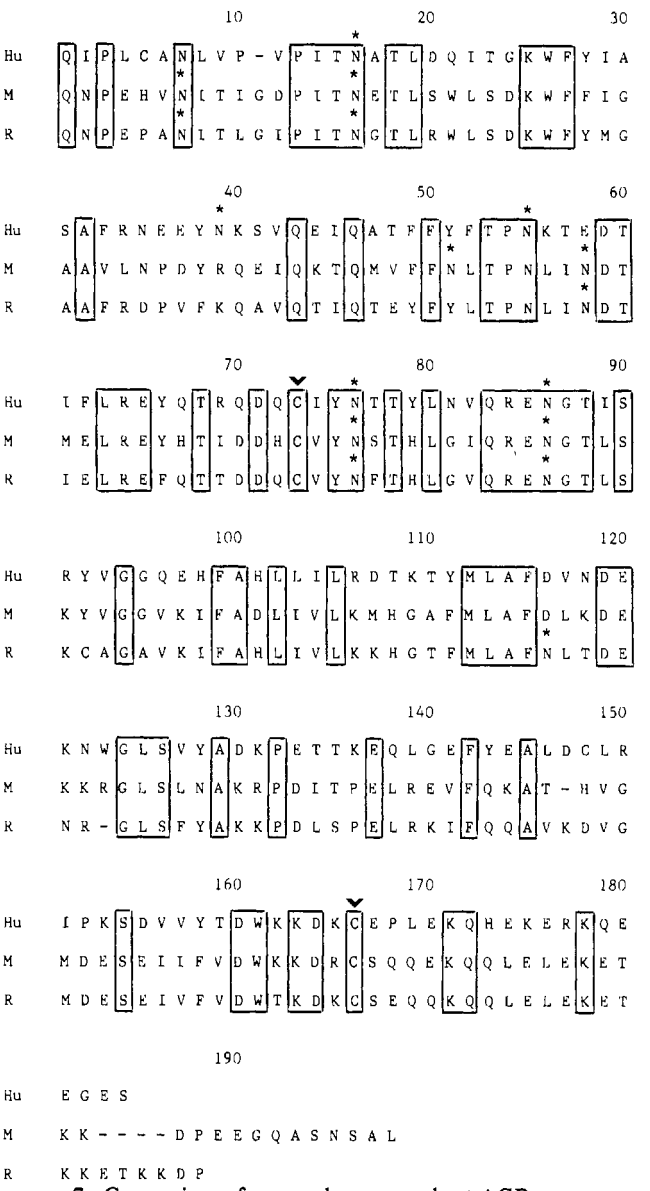

FIGURE 7: Comparison of mouse, human, and rat AGP sequences. The amino acid sequence of mouse AGP is shown compared to the sequences of human and rat AGP. Amino acids that are identical in all three species are enclosed in boxes. Arrows indicate the positions of the cysteine residues. Asterisks mark the position of putative glycosylation sites as indicated by the presence of the sequence Asn-X-$^{Thr}_{Ser}$ (N-X-$^{T}_{S}$).

contrast, sequence analysis of the same region in the mouse AGP-1 gene has shown potential glucocorticoid receptor binding sites at −127 to −122 (AGAACA) and at −94 to −79. The ability of this region of the mouse gene to mediate a response to glucocorticoids must be demonstrated. Furthermore, since there are multiple AGP genes in the mouse, analyses must be done to determine whether all genes respond to glucocorticoids and whether the sequences that mediate this response are conserved.

Dente et al. (1985) have observed a region of striking DNA sequence homology between three otherwise unrelated human acute-phase proteins including one of the human AGP genes. We have searched the mouse AGP-1 gene as well as the published rat AGP gene sequences (Reinke & Feigelson, 1985; Liao et al., 1985) for this same sequence and found a region of very close homology in both rodent species in an exactly analogous position, i.e., between the TATAA box and the initiator methionine codon. Because of its location partly within the 5′ untranslated portion of the primary mRNA

transcript, it has been suggested (Dente et al., 1985) that this element might be involved in the posttranscriptional regulatory mechanism involving glucocorticoid-regulated RNA processing (Vannice et al., 1984). In light of the most recent evidence supporting the transcriptional regulation of the AGP gene in vivo, (Kulkarni et al., 1985), we favor a model in which this element may modulate the transcription of the gene in a positive or negative manner.

ACKNOWLEDGMENTS

We thank Dr. David Konkel for his critical review of the manuscript and his suggestions and Glenda M. Romero for her assistance in the preparation of the manuscript.

REFERENCES

Alsip, G., & Konkel, D. A. (1984) *BRL Focus 6(13)*, 12.
Baumann, H., & Maquat, L. E. (1986) *Mol. Cell. Biol. 6*, 2551–2561.
Baumann, H., Firestone, G. L., Burgess, T. L., Gross, K. W., Yamamato, K. R., & Held, W. A. (1983a) *J. Biol. Chem. 258*, 563–570.
Baumann, H., Jahreis, G. P., & Gaines, K. C. (1983b) *J. Cell Biol. 97*, 866–876.
Baumann, H., Held, W. A., & Berger, F. G. (1984a) *J. Biol. Chem. 259*, 566–573.
Baumann, H., Jahreis, G. P., Sauder, D. N., & Koj, A. (1984b) *J. Biol. Chem. 259*, 7331–7342.
Benton, W. D., & Davis, R. W. (1977) *Science (Washington, D.C.) 196*, 180–182.
Biggin, M D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.
Breathnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.
Cooper, R., & Papaconstantinou, J. (1986) *J. Biol. Chem. 261*, 1849–1853.
Cooper, R., Herzog, C. E., Li, M.-L., Zapisek, W. F., Hoyt, P. R., Ratrie, H., III, & Papaconstantinou, J. (1984) *Nucleic Acids Res. 12*, 6575–6586.
Dente, L., Ciliberto, G., & Cortese, R. (1985) *Nucleic Acids Res. 13*, 3941–3952.
Enquist, L. W., Madden, M. J., Schiop-Stansly, P., & Vande Woude, G. F. (1979) *Science (Washington, D.C.) 203*, 541–544.
Eylar, E. H. (1966) *J. Theor. Biol. 10*, 89–113.
Favaloro, J., Triesman, R., & Kamen, R. (1980) *Methods Enzymol. 65*, 718–749.
Fowlkes, D. M., Mullis, N. T., Comeau, C. M., & Crabtree, G. R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2313–2316.
Friedman, M. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5421–5424.
Hamada, H., & Kakuraga, T. (1982) *Nature (London) 298*, 396–398.
Hamada, H., Seidman, M., Howard, B. H., & Gorman, C. M. (1984) *Mol. Cell. Biol. 4*, 2622–2630.
Hentschel, C. C. (1982) *Nature (London) 295*, 714–716.
Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M., & Beato, M. (1984) *Nature (London) 308*, 513–519.
Khoury, G., & Gruss, P. (1983) *Cell (Cambridge, Mass.) 33*, 313–314.
Koj, A. (1974) *Structure and Function of Plasma Proteins* (Allison, A. C., Ed.) Vol. 1, pp 73–125, Plenum, London.
Kulkarni, A. B., Reinke, R., & Feigelson, P. (1985) *J. Biol. Chem. 260*, 15386–15389.
Lei, K.-J., Liu, T., Zon, G., Soravia, E., Liu, T.-Y., & Goldman, N. D. (1985) *J. Biol. Chem. 260*, 13377–13383.
Liao, Y.-C., Taylor, J. M., Vannice, J. L., Clawson, G. A., & Smuckler, E. A. (1985) *Mol. Cell. Biol. 5*, 3634–3639.
Long, G. L., Chandra, T., Woo, S. L. C., Davie, E. W., & Kurachi, K. (1984) *Biochemistry 23*, 4828–4837.
Maeda, H. (1985) *J. Biol. Chem. 260*, 6698–6709.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Marx, J. L. (1985) *Science (Washington, D.C.) 230*, 794–796.
Miesfield, R., Krystal, M., & Arnheim, N. (1981) *Nucleic Acids Res. 9*, 5931–5947.
Nagashima, M., Urban, J., & Schreiber, G. (1981) *J. Biol. Chem. 256*, 2091–2093.
Nishioka, Y., & Leder, P. (1980) *J. Biol. Chem. 255*, 3691–3694.
Norgard, M. V., Emigholz, K., & Monahan, J. J. (1979) *J. Bacteriol. 138*, 270–272.
Northemann, W., Andus, T., Gross, V., Nagashima, M., Schreiber, G., & Heinrich, P. C. (1983) *FEBS Lett. 161*, 319–322.
Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F., & Sharp, P. A. (1984) *Science (Washington, D.C.) 225*, 898–903.
Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., & Surrey, S. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4298–4302.
Reinke, R., & Feigelson, P. (1985) *J. Biol. Chem. 260*, 4397–4403.
Ricca, G. A., & Taylor, J. M. (1981) *J. Biol. Chem. 256*, 11199–11202.
Ricca, G. A., McLean, J. W., & Taylor, J. M. (1982) *Ann. N.Y. Acad. Sci. 389*, 88–103.
Rich, A., Nordheim, A., & Wang, A. H.-J. (1984) *Annu. Rev. Biochem. 53*, 791–846.
Ritchie, D. G., & Fuller, G. M. (1983) *Ann. N.Y. Acad. Sci. 408*, 490–500.
Ritchie, D. G., Levy, B. A., Adams, M. A., & Fuller, G. M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1530–1534.
Ruskin, B., Krainer, A. R., Maniatis, T., & Green, M. R. (1984) *Cell (Cambridge, Mass.) 38*, 317–331.
Sanger, F., Nicklen, S., & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schmid, K. (1975) *The Plasma Proteins* (Putnam, F., Ed.) Vol. 1, pp 184–228, Academic, New York.
Scott, R. W., & Tilghman, S. M. (1983) *Mol. Cell. Biol. 3*, 1295–1309.
Searle, P. F., Stuart, G. W., & Palmiter, R. D. (1985) *Mol. Cell. Biol. 5*, 1480–1489.
Shen, S., Slightom, J. L., & Smithies, O. (1981) *Cell (Cambridge, Mass.) 26*, 191–203.
Slater, E. P., Rabenau, O., Karin, M., Baxter, J. D., & Beato, M. (1985) *Mol. Cell. Biol. 5*, 2984–2992.
Vannice, J., Taylor, J. M., & Ringold, G. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4241–4245.
Walmsley, R. M., Szostak, J. W., & Petes, T. D. (1983) *Nature (London) 302*, 84–86.
Woo, P., Korenberg, J. R., & Whitehead, A. S. (1985) *J. Biol. Chem. 260*, 13384–13388.
Yannisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103–119.
Zeitlin, S., & Efstratiadis, A. (1984) *Cell (Cambridge, Mass.) 39*, 589–602.