Turnell, R. W., Kaiser, N., Milholland, R. J., & Rosen, F. (1974) *J. Biol. Chem. 249*, 1133-1138.
Van Bohemen, C. G., & Rousseau, G. G. (1982) *FEBS Lett. 143*, 21-25.
Welshons, W. V., Krummel, B. M., & Gorski, J. (1985) *Endocrinology (Baltimore) 117*, 2140-2147.
Westphal, H. M., Moldenhauer, G., & Beato, M. (1982) *EMBO J. 1*, 1467-1471.
Willman, T., & Beato, M. (1986) *Nature (London) 324*, 688-691.
Wira, C. R., & Munck, A. (1974) *J. Biol. Chem. 269*, 5328-5336.

Structure and Acute-Phase Regulation of the Rat α_2-Macroglobulin Gene†

Wolfgang Northemann,‡ Brian R. Shiels,‡,§ Todd A. Braciak,‡ Richard W. Hanson,‖ Peter C. Heinrich,⊥ and Georg H. Fey*,‡

Department of Immunology, Research Institute of Scripps Clinic, La Jolla, California 92037, Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, and Institute of Biochemistry, University of Freiburg, Freiburg, West Germany

Received April 1, 1988; Revised Manuscript Received July 28, 1988

ABSTRACT: Seven genomic DNA clones representing the rat α_2-macroglobulin gene were isolated and characterized. The cloned sequence covered the entire gene (48 kilobases) plus 2 kilobases of 3′- and 13.7 kilobases of 5′-flanking sequences. A restriction cleavage map of the gene was produced, and the restriction cleavage pattern of genomic DNA suggested that the α_2-macroglobulin gene is a single-copy gene. A 7.7-kilobase fragment from the 5′-terminal region and a 250 base pair fragment from the 3′-terminal region of the gene were sequenced, and the 3′ end of the gene was mapped. The sequenced 5′-terminal fragment contained 4.5 kilobases of 5′-flanking sequences plus the first three exons and two introns of the gene. Two transcription start sites, a minor and a major site, located 65 nucleotides apart, were defined by primer extension, S1 mapping, and RNaseH experiments. During an acute-phase response, transcription from both sites was induced in the liver, and over 90% of the transcripts originated from the major site. Very high concentrations of α_2-macroglobulin mRNA originating from both start sites were also found in the uterus but not in the liver of pregnant females. A glucocorticoid response element (GRE), a conserved consensus sequence for a potential glucocorticoid receptor DNA binding site, was found by computer search in the promoter-proximal 5′-flanking region of the α_2-macroglobulin gene.

The α_2-macroglobulin (α_2M)[1] is a plasma glycoprotein of high molecular weight (approximate M_r 180 000) which is mainly synthesized in the parenchymal cells of the liver but also in a number of other cell types. During acute and chronic inflammations occurring in response to tissue damage and infections, α_2M concentrations in the plasma are dramatically increased in rats, and α_2M is the major acute-phase protein in rats (Gordon, 1976; Gordon & Koj, 1985; Schreiber, 1987; Lonberg-Holm et al., 1987; Northemann et al., 1985; Hayashida et al., 1985; Gehring et al., 1987). α_2M is a member of the family of rat α-macroglobulins, which includes also α_1-macroglobulin (α_1M) and several species of α_1-inhibitor III (α_1I3). All α-macroglobulins are proteinase inhibitors of high molecular weight. All members of this family contain an internal thiolester bond, which is important for their function, and all are structurally related (Lonberg-Holm et al., 1987; Sottrup-Jensen, 1987; Gehring et al., 1987; Braciak et al., 1988; Gauthier & Ohlsson, 1978; Esnard & Gauthier, 1980; Schaeufele & Koo, 1982). The α-macroglobulins are further structurally related to complement components C3, C4, and C5, and the genes for all family members have evolved from a common ancestor (Sottrup-Jensen et al., 1985; Sottrup-Jensen, 1987). cDNA clones for the major rat α-macroglobulins have been isolated and characterized and were used to establish that the changes of macroglobulin concentrations occurring in the plasma during an acute-phase response are preceded by corresponding changes of their mRNA concentrations in the liver. These changes in mRNA concentrations are caused at least in part by increased transcription of the α_2M gene (Northemann et al., 1985, 198b; Hayashida et al., 1985; Birch & Schreiber, 1986; Gehring et al., 1987). Posttranscriptional alterations in the stability of α_2M mRNA may also participate in the overall changes of α_2M mRNA concentration that occur during an acute-phase response, but they have not yet been demonstrated directly. Transcription of the α_2M gene is increased at least 5-8-fold

† This work was supported by Grant AI22166 from the National Institute of Allergy and Infectious Diseases (to G.H.F.), by a seed grant from the Liver Association (to G.H.F.), and by short-term awards from the North Atlantic Treaty Organization, the Boehringer Ingelheim Fonds (to W.N.), and the Fulbright Commission (to B.R.S.). Additional support was provided by the Deutsche Forschungsgemeinschaft (Grant He906 to P.C.H.) and by Grants AM21859 and DK25541 from the National Institutes of Health (to R.W.H.). This is Publication No. 4951-IMM from the Department of Immunology, Research Institute of Scripps Clinic, La Jolla, CA.

* Author to whom correspondence should be addressed at the Department of Immunology, IMM-14, Research Institute of Scripps Clinic, 10666 North Torrey Pines Rd., La Jolla, CA 92037.
‡ Research Institute of Scripps Clinic.
§ Present address: Welcome Unit of Molecular Parasitology, Department of Veterinary Parasitology, Bearsden Road, Glasgow G61 1QH, Scotland.
‖ Case Western Reserve University.
⊥ University of Freiburg.

[1] Abbreviations: α_2M, α_2-macroglobulin; α_1M, α_1-macroglobulin; α_1I3, α_1-inhibitor III; AGP, α_1-acid glycoprotein; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); bp, base pair(s); kb, kilobase pair(s); HSF, hepatocyte stimulating factor; IL6, interleukin 6; GRE, glucocorticoid responsive element.

0006-2960/88/0427-9194$01.50/0

during an acute-phase response, and probably more (Northemann et al., 1985, 1988b; Birch & Schreiber, 1986). Therefore, together with the other α-macroglobulin genes, the α_2M gene presents an excellent model for biochemical and genetic studies of control mechanisms of transcription and probably also of posttranscriptional RNA processing.

The α_2M gene is regulated coordinately by two types of hormonal signals: glucocorticoids and hepatocyte stimulating factor (HSF), also called interleukin 6 (IL6) [Northemann et al., 1988a; Baumann et al., 1987; Gauldie et al., 1987; for review, see Fey and Fuller (1987)]. It is our long-range objective to determine how these two hormones modulate transcription of the α_2M gene. As a first step toward an analysis of the cis- and trans-acting transcription control elements of this gene, we have isolated and characterized the rat α_2M gene. Here we report the restriction map of the gene, the localization of its 3′ and 5′ ends, and the sequence of 4.5 kilobases of its 5′-flanking region as well as of its first three exons and two introns. We have found that this gene is transcribed from two adjacent start sites and have investigated whether these start sites are differentially utilized in different cell types producing α_2M. No differential utilization of both transcription start sites was detected in the cell types examined. In the course of these studies we observed that this gene is strongly transcribed in the uterus but not in the liver of pregnant females. Therefore, the gene is probably controlled by different hormonal signals in the uterus and in the liver.

This observation has prompted us to search for the consensus sequences for hormone responsive control elements in the 5′-terminal region of the α_2M gene.

Experimental Procedures

Animals and Materials. Male Fisher 344 rats were from Simonsen Laboratories, Gilroy, CA. Synthetic oligonucleotides were produced on an Applied Biosystems Model 380A DNA synthesizer, using β-cyanoethyl phosphoramidite chemistry.

Isolation and Characterization of Genomic DNA Clones. A genomic DNA library constructed from DNA fragments partially digested with *Hae*III (Sargent et al., 1981) was screened with a 600-bp cDNA probe representing the 3′ end of the α_2M mRNA (Northemann et al., 1985), leading to the isolation of λRA2MG-73. A restriction map of this cloned DNA was produced, and a terminal *Eco*RI restriction fragment from this clone was used to rescreen the library. In successive rounds of rescreening the same library, clones λRA2MG-73, -12, -22, -31, and -53 were isolated. The area covered by clone λRA2MG-44, isolated from the *Eco*RI partial digest library (Sargent et al., 1979), was never covered by a clone from the *Hae*III partial library. Similarly, we failed to isolate clones covering the 5′-flanking region beyond the region spanned by clone λRA2MG-53 from both the *Eco*RI and the *Hae*III partial libraries. Therefore clone λRA2MG-63 was isolated from a third library, constructed from DNA partially digested with *Sau*IIIA in the vector λ EMBL 3 (Schmid et al., 1982) by screening with a 5′-terminal cDNA fragment from a full-length cDNA clone (Gehring et al., 1987). In this library the insert DNA fragments were coupled to the vector with *Sal*I linkers. DNA was prepared from all of these clones, and restriction maps were produced by single and double digestions with the enzymes *Bam*H1, *Eco*RI, *Hin*dIII, and *Sst*I. Insert DNA fragments were subcloned into the plasmid vector pBR328 (Soberon et al., 1980), and subsubclones were prepared for sequencing from selected regions in the single-stranded phage vector M13mp8 (Bankier & Barrell, 1983).

Southern Blot Analysis of Genomic DNA. Genomic DNA from the livers of male rats was isolated and digested to completion with *Eco*RI. In parallel, DNA from selected genomic phages was digested with *Eco*RI, and all fragments were separated by electrophoresis in a 1% agarose gel. Transfer to Bio-Rad Zeta Probe nylon membrane by capillary blotting, prehybridization, hybridization, and wash procedures followed standard protocols (Maniatis et al., 1982).

DNA Sequence Analysis. A 2.02-kb *Sal*I/*Hin*dIII fragment, a 1.36-kb *Hin*dIII/*Hin*dIII fragment, and a 1.60-kb *Hin*dIII/*Sal*I (artificial) fragment from clone λRA2MG-63 and a 1.88-kb *Eco*RI/*Pst*I fragment, as well as a 1.52-kb *Pst*I/*Bam*HI fragment, from clone λRA2MG-53 were randomly sheared by sonication and subcloned into the M13mp8 vector (Bankier & Barrell, 1983). DNA sequencing was performed according to the dideoxynucleotide sequencing technique (Sanger et al., 1977). DNA sequences were analyzed with Staden's DB-system and ANALYSEQ programs (Staden, 1986) and the University of Wisconsin Genetics Computer Group Programs (Devereux et al., 1984) on a VAX 11/750 computer. Sequencing of random DNA clones was continued until each character of the final sequence was covered at least once, preferably twice, on each strand. On the average, each nucleotide was read from 9 to 10 independently sequenced M13 clones, resulting in a database of 73 203 characters for the 7.7-kb sequence of Figure 5.

Preparation of RNA, Size Enrichment, and Northern Blot Analysis. These experiments were performed as previously described (Gehring et al., 1987; Northemann et al., 1985; Braciak et al., 1988; Shiels et al., 1987). RNA from H35 hepatoma cells (Baumann et al., 1987) was prepared with guanidinium thiocyanate and sedimented through a cesium chloride cushion (Turpen & Griffith, 1986).

S1 Nuclease Mapping. One hundred fifty nanograms of a 928-bp *Pst*I/*Hin*dIII fragment was prepared and labeled at a single end as described in the legend of Figure 6A. One aliquot of this material was annealed with 3 μg of size-enriched polyadenylated RNA from livers of experimentally inflamed rats and another with 5 μg of size-enriched poly(A+) RNA from control rats. Annealing was performed in a 50-μL reaction volume, containing 80% formamide, 40 mM PIPES, pH 6.4, 0.4 M sodium chloride, and 1 mM EDTA, at 52 °C for 4 h. The reaction was chilled on ice, diluted with 450 μL of ice-cold S1 buffer containing 280 mM sodium chloride, 4.5 mM zinc acetate, and 30 mM sodium acetate, pH 4.4, and incubated with 400 units of S1 nuclease for 45 min at 37 °C. The nucleic acids were extracted with phenol and precipitated with ethanol. The precipitate was redissolved in 8 μL of 10 mM Tris-HCl and 0.1 mM EDTA, pH 8.0, and an aliquot was electrophoresed in a 6% polyacrylamide 6 M urea-sequencing gel (Bankier & Barrell, 1983). To prepare internally labeled, single-stranded DNA, the 0.5l-kb *Eco*RI/*Sal*I fragment of λRA2MG-63 containing exon 1 and the 1.55-kb *Pst*I/*Bam*HI fragment of clone λRA2MG-53 containing exons 2 and 3 were subcloned into the phage vector M13mp8. Internally labeled phage M13 DNA was prepared as described (Northemann et al., 1988b) and used for S1 nuclease mapping. Seven micrograms of size-enriched polyadenylated RNA extracted from the livers of rats 18 h after initiation of the acute-phase response was annealed with 0.5 μg of internally labeled M13 phage DNA, containing exon 1 and exons 2 plus 3, respectively. Treatment with S1 nuclease and electrophoretical analysis was carried out as described (Northemann et al., 1988b).

Primer Extension Analysis. Approximately 150 ng of a *Hpa*II/*Bam*HI cDNA fragment, representing nucleotides

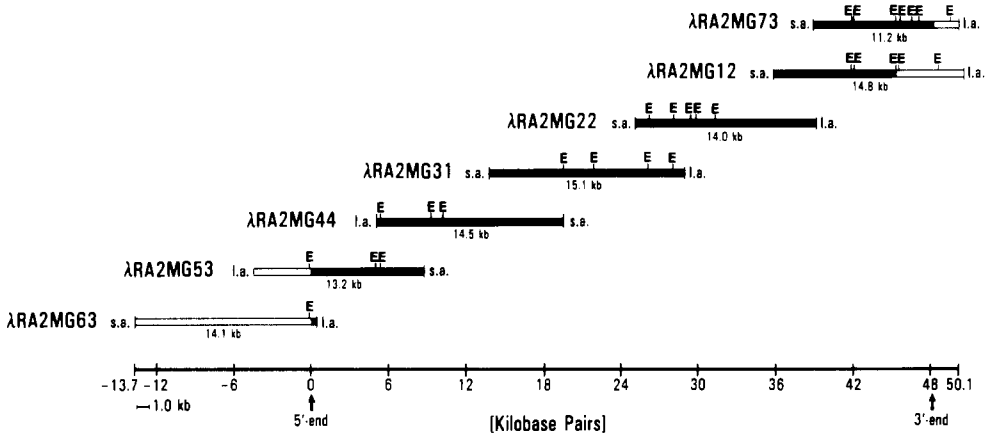

FIGURE 1: Genomic DNA clones containing the rat α_2-macroglobulin gene. The inserts of seven overlapping genomic DNA clones, contained in phage λ, are represented by boxes. Closed parts of the boxes represent the transcription unit and open parts the flanking regions of the gene. The number under each insert gives its length in kb. E, restriction cleavage sites for the enzyme *Eco*RI; s.a. and l.a., attachment sites of the short arm and long arm of the λ vector. Arrows designate the 5′ and 3′ ends of the transcription unit. Phages λRA2MG-12, -22, -31, -53, and -73 were from the *Hae*III partial digest gene library (Sargent et al., 1981), phage λRA2MG-44 was from the *Eco*RI partial digest library (Sargent et al., 1979), and phage λRA2MG-63 was from a third library, constructed from partially *Sau*IIIA-digested DNA with *Sal*I linkers in the vector λ EMBL 3 (Schmid et al., 1982). No clones overlapping the insert of clone 44 were detected in the *Hae*III library and no clones overlapping the insert of clone 63 in either the *Hae*III or *Eco*RI partial digest libraries.

81–576 of the published cDNA sequence (Gehring et al., 1987), was used as a primer and annealed with 3 μg of size-enriched polyadenylated liver RNA from control rats and experimentally inflamed rats, respectively. Strand separation of the primer was achieved by heating for 10 min at 67 °C in 90% formamide in the absence of salts. Then salts were added, and annealing was performed in a 50-μL reaction volume, containing 80% formamide, 10 mM PIPES, pH 6.4, 0.4 M sodium chloride, and 1 mM EDTA, at 55 °C for 4 h. The mixture was cooled on ice and diluted with 50 μL of the same buffer. Nucleic acids were precipitated with ethanol and resuspended in 25 μL of primer extension reaction buffer, containing 25 mM Tris-HCl, pH 8.3, 50 mM potassium chloride, 0.5 mM each of dCTP, TTP, and dGTP, 5 μM dATP, 50 μCi of [^{32}P]dATP (sp act. 3000 Ci/mmol), 50 ng/μL actinomycin D, 1 mM dithiothreitol (DTT), and 5 mM magnesium chloride. The extension reaction was performed with 26 units of reverse transcriptase at 43 °C for 90 min. Subsequently, dATP was added to 0.5 mM together with another 26 units of reverse transcriptase, and the incubation was continued for another 30 min at 43 °C. The RNA template was then degraded with sodium hydroxide, the reaction mixture was neutralized, and nucleic acids were precipitated with ethanol. The precipitate was redissolved in 10 mM Tris-HCl and 0.1 mM EDTA, pH 8.0, and an aliquot was applied to a sequencing gel (Bankier & Barrell, 1983).

RNaseH Mapping. Size-enriched, polyadenylated liver RNA (10 μg) from normal and acute-phase rats, respectively, was annealed with 600 ng of the synthetic oligonucleotide A, representing nucleotides 289–328 of the published α_2M cDNA sequence (Gehring et al., 1987). For primary rat hepatocytes and H35 rat hepatoma cells, the polyadenylated RNA (20 μg) was not size enriched. Annealing was performed in a 100-μL reaction volume, containing 80% formamide, 10 mM PIPES, pH 6.4, 0.4 M sodium chloride, and 1 mM EDTA, at 50 °C for 90 min. The annealing mix was diluted with 100 μL of the same buffer, and nucleic acids were precipitated with ethanol. The precipitate was resuspended in 100 μL of 20 mM Tris-HCl, pH 7.4, 100 mM potassium chloride, 10 mM magnesium chloride, and 0.1 mM DTT, containing 8 units of RNaseH. After incubation for 60 min at 37 °C, the nucleic acids were extracted with phenol and precipitated with ethanol. The precipitate was redissolved in 12 μL of 10 mM Tris-HCl, pH 8.0, and 0.1 mM EDTA, and a 5-μL aliquot was applied to a sequencing gel (Bankier & Barrell, 1983). The RNA was transferred to Pall Biodyne nylon membrane by electroblotting using a Bio-Rad transblot cell and immobilized by brief illumination with ultraviolet light as described (Gehring et al., 1987; Shiels et al., 1987). Hybridization of this blot with radiolabeled oligonucleotides B and F was performed as described (Northemann et al., 1988b).

RESULTS

A set of seven overlapping rat genomic DNA clones covering the entire rat α_2M gene was isolated by reiterated screening of three different genomic DNA libraries with cloned cDNA probes as described under Experimental Procedures. Restriction enzyme cleavage maps were produced for the enzymes *Eco*RI, *Bam*HI, *Hin*dIII, and *Sst*I, and the *Eco*RI map is shown in Figure 1. A Southern blot experiment using nick-translated, full-length cDNA probes (Gehring et al., 1987) under conditions known to eliminate cross-hybridization with other closely related members of the rat α-macroglobulin family confirmed that all genomic clones of this series hybridized with α_2M cDNA. Thus, these clones contain only α_2M gene sequences but not sequences of closely related other α-macroglobulin genes. Total rat genomic DNA was digested with *Eco*RI and hybridized with a mixture of short DNA probes representing only extreme 5′- and 3′-terminal coding regions of the gene (Figure 2). The same restriction fragments of 5.2, 3.25, and 2.40 kb were observed in the genomic DNA and in the mixture of genomic DNA clones, with no additional fragments visible in the genomic DNA. In addition, during the multiple rounds of rescreening, no variant clones with different restriction patterns were isolated from these genomic libraries. We thus conclude that the rat genome carries only a single α_2M gene locus, comparable with the single α_2M locus observed in the human genome (Kan et al., 1985). No restriction polymorphic alleles were discovered between the inbred strains used for the construction of the genomic libraries and for the preparation of genomic DNA used for the Southern blot in Figure 2.

The 3′ end of the gene was located by comparison of genomic DNA sequences determined from parts of the DNA insert of clone λRA2MG-73 with published α_2M cDNA sequences (Gehring et al., 1987). A perfect match was found

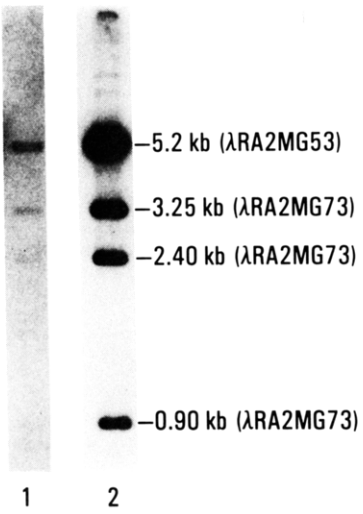

FIGURE 2: Southern blot analysis of rat genomic DNA. (Track 1) Twenty micrograms of rat genomic DNA was digested with *Eco*RI. The fragments were separated by electrophoresis and blotted to Zeta Probe membrane, as described under Experimental Procedures. (Track 2) A mixture of 0.2 μg of DNA from each of the phages λRA2MG-53 and -73 was digested with *Eco*RI and blotted. Hybridization was performed with a mixture of two nick-translated cDNA fragments, representing the 3′- and 5′-terminal portions of the gene in order to reduce the complexity of the band pattern. The 3′-end probe was a 450-bp *Pst*I fragment; the 5′-end probe was the 495-bp *Hpa*II/*Bam*HI fragment of cDNA clone pRLA2M-29 (Gehring et al., 1987). The sizes of the hybridizing fragments (in kb) are given on the right; in parentheses, the genomic DNA clones containing these fragments are given. Phage λ DNA fragments digested with *Hin*dIII were used as size markers.

```
5'   cttaattttctcttctttccactaccacagATTATGGAAATGCCTCAGGA
                                   ATTATGGAAATGCCTCAGGA

     CGCAGTGAATAAGAAGTGTTTCGCCAGAGCCCTGACCTCAGGACTTCCCA
     CGCAGTGAATAAGAAGTCTTTCGCCAGAGCCCTGACCTCAGGACTTCCCA

     AGAAAAACAGTGTATTTGTATTTCCAGAGATTTGATCAATAAACCATTTT
     AGAAAAACAGTGTATTTGTATTTCCAGAGATTTGATCAATAAACCATTTT

     TTTCATATCTAcccatcaccattccctcattaattcagtcatcatttttt
     TTTCATATCTA-(A)n

     tacattttcatgcttccacatctagtactcaataaaaactgatgaaaatc   3'
```

FIGURE 3: Determination of the 3′ end of the α_2M gene. A genomic DNA fragment from clone λRA2MG-73 was randomly subcloned into M13 vectors and sequenced (top row), and the sequence was compared with the published cDNA sequence (Gehring et al., 1987) (bottom row). Exon sequences are in upper-case letters; intron sequences are in lower-case letters. The position of the poly(A) tail in the mRNA is indicated by (A)n. A conventional polyadenylation signal AATAAA precedes the polyadenylation site at the standard distance of 17 nucleotides.

for 131 nucleotides (Figure 3). At the point of divergence, the cDNA sequence continued with a homopolymeric stretch of A residues, which probably represent the start of the poly(A) tail. A typical polyadenylation signal, AATAAA, was found 17 nucleotides upstream of the point of sequence diversion, and we concluded that this point represents the 3′ end of the gene.

The sequence of a contiguous 7.7-kb stretch of cloned genomic DNA including 4.5 kb of 5′-flanking region and the first three exons and two introns of the gene was determined (Figure 4), and a schematic representation of the exon–intron block structure of the sequenced region including a restriction map is given in Figure 5.

The 5′ ends of the α_2M RNA transcripts were mapped on the gene by three independent procedures: S1 nuclease mapping, primer extension, and RNaseH mapping. For S1 nuclease mapping we have used both double-stranded DNA which was labeled at a single end and internally labeled, single-stranded genomic DNA (Figure 6, panels A and B). With genomic DNA, which was labeled at a single end, two groups of S1-resistant RNA/DNA hybrids were observed (Figure 6A, tracks 1 and 2). One quintuplet of S1-protected fragments defined the 5′ ends of one size class of RNA molecules, while a second quadruplet of weaker intensity, located 65 nucleotides in the 5′ direction, defined the 5′ ends of a second RNA size class. Both of these S1-protected DNA fragments were observed only when RNA from acute-phase liver was used (Figure 6A, tracks 1 and 2), but not with RNA from normal rat liver (Figure 6A, track 3). Therefore, the concentration of both RNA species is greatly stimulated in the livers of rats during an acute-phase reaction, but over 90% of the RNA molecules originated from the major start site, which is located closer to the coding portion of the gene. From a combination of data from S1 nuclease mapping, primer extension, and RNaseH mapping experiments (see below), we have concluded that the A residue indicated as M(87) in Figure 6A corresponds to the major transcription start site. From similar experiments we deduced that the 5′ end of the longer, less abundant RNA transcripts maps at the G residue labeled m(152) in Figure 6A. In the coordinates of the sequence given in Figure 4, these two residues are located at positions 4464 and 4529, respectively. Both sites are preceded at distances of 18 and 15 nucleotides, respectively, by pentanucleotides resembling a TATA–box consensus sequence (Corden et al., 1980). The pentanucleotide preceding the major transcription start site ($TATA_M$) is ATAAA, while the element for the minor site ($TATA_m$) is AAAAA.

S1 mapping with internally labeled, single-stranded genomic DNA confirmed the map position of the major transcription start site but was not sensitive enough to clearly reveal the map position of the minor start site. A faint band of appropriate size for the minor RNA size class was detected, but it was located in a region of high background, which precluded its further evaluation (Figure 6B, track 1). However, this technique resulted in exactly the same values for the lengths of exons 2 and 3 as determined by a comparison of cDNA and genomic DNA sequences (Figure 6B, track 2). Thus, the length of exon 1 was reliably determined at 164 nucleotides (Figure 6B, lane 1).

Primer extension experiments were performed with a cloned cDNA fragment 495 nucleotides in length as a specific primer (data not shown) and led to the location of the 5′ end of the major RNA class at the C residue four nucleotides downstream of the site M(87) as defined above, corresponding to nucleotide 4533 in Figure 5. These two results are in good agreement, considering the error limit for the size determination of the primer extension products of ±5 nucleotides.

A third independent experiment was performed to map the α_2M RNA 5′ ends (Figure 7). A synthetic oligonucleotide (called oligonucleotide A) corresponding to map positions 289–328 of the published α_2M cDNA sequence [Gehring et al. (1987), see Figure 7] was annealed with polyadenylated, size-enriched RNA from the livers of both normal and acutely inflamed rats. The reaction mixtures were treated with RNaseH, which selectively degrades RNA in RNA/DNA heteroduplex conformation but leaves single-stranded RNA intact. The RNaseH-resistant RNA fragments were electrophoretically separated in a polyacrylamide gel and transferred to a nylon membrane. Duplicate blots were hybridized separately with two other radiolabeled oligonucleotides: oligonucleotide B representing a sequence 3′ of the major transcription start site M(87) [nucleotides 30–50 of the published α_2M cDNA sequence, Gehring et al. (1987), Figure 7] and

```
   1 GTCGACAACT TACCTTTCTC GTTTTGAAGG TAGTTAATGT GTAAACAGCT GAAATGTCCT TATCATCACG AGGGACAGAA GATCTGTTTC CCATGTTCTC
 101 TTGGCTTAGG CGACGTGAGC TAGTGCACAC CAAGTACAAT CACAGAACTT TCTCTTTCCT GTAAGTCATT TTGCAATGTG AGGGGGAAGT GTTTTCTTAC
 201 TAATTTTTAT TAAGGCCTAT TTGTCTCCTA ATCTTTTTTT TCTCAACTCC CAAATATGCC CTCGTATTCT TATTGCCACG TGCTCCTTCT GTGTTCTGCA
 301 CCCATGAACT GCCTTCCTTA TTTTTTTACT TTCAGTTACC CTATCTCATA TTTCTGACTG TAACACTCGC TAAGTTCTTC AAAGTTGTAT AAAACCCTCC
 401 ACATCGTGTG TTCCCTTTCC CATGAGGGAT GGTCTATTCA GGTAAGGAAA CTCACATTTT TTTTTCTCCC TGGAAATATT TCTGGACACG GTGCCTTTGG
 501 CCAAACCCAC AGAAGATACA AATAATCTCT TCAACCACTG ATGTAAAATA TAAAGATTGA TTTTTCTCTA ATGTGATAAG GTTTACAGGT TAAAGGTCAC
 601 TGCTGCAGCG TCTGGTCCAC ATTTAGCTCC GTGAGCTTAT AGCCCATATG CCAGGCTTCA AACTTGCTGG CAGGACAGCC GAGCCCTGGA GCAGATAGCT
 701 TGGCTGGCCA GGACCACATG GTTCAGTCAG AACTGAGGTT CTGGTGACCC CGCAACTGAC TGCCCTGCCC ATTATCAACT GTTCCTTCTC AAACTCCCCC
 801 ACCCCAAGTG AAATAGGTTT GAAAAACTTT TCCTCCTTAT GAATATGTGA GGAATGGGAT ACATGAGGAG AGAATGTTCC TCCACCAGGG GTCATTAATT
 901 AATTGCTCCA CACATTATTA TTATTATTAT TATTATTATT ATTATTATTA TTATTGTTGT TGTTGTTGTT GTTGTTGTTG TTATTGTTGT TCCACTCATC
1001 GACACAGTGT GTGCAGATCA ATTTCAACTC TCCTTTTTTC TCTTACAGTT TCCCAAGCTC AAGTAATCAA GATTATACTT CCAACTGACT TGAATAATTT
1101 CCATCTATGC ATTAAAAGCT GAAGCAGTTG CCTTTCTGTG TCTGGCCTGT TTAACTTAAC ACAGCATTCT GCCATCCTTT CATTTATTGC ACACGACAGA
1201 ACATAACTGA ATAGTGTTTC ATCGCGATAC ACATCTGAAC ACATCTGCTC ATCCTGTCCT CAGAGTCTCC CTCTGAAATG ATTGCTGCAC ACACATTTGA
1301 AATTCTTATC TGTGGCTGTG ATAACATCCT GACAAGAGCA ACTTAAGGGC TCAAGGTTTA GGGTACATCG TGTTAGGAAA TCCAAAGCCA CAGGACCATG
1401 AAGCACATGG ACACATCGCA CCCACAGTGG AAACAGACCT TGGGTTTCCT CCTGCATTTT CACACTGTGT AAGGAATGGT CCCACCCAAG GCAGGTAGGT
1501 CTTCCAACAT CAGAACCTCA GCTATGGTAG CCAAGACATC CACAGGGATG TGTAGAGTCC CATTTTCTAG GATAGTCTAG ATTCTGCCAA GCTGACAACT
1601 CACTGTAGGA ATCATGGCAG CTCAGCCTTT CCCCTTTCAT TGGTTGCTCC TCCCCCATTT CATTTGGTTG CTCCTCCCTA CTACGTTGGT TGCTCCTCCT
1701 CCTACGTTGG TTGCTCCTCC CCCTACGTTG GTTGCTCCTC CCCCAACATG GTTGCTCCTC CCCCTTACAT TGGTTGCCCC TCCCCCTTAC GTTGACTGCT
1801 CCCCTTACAT TGGGTACTCC TTCCCCACTT ACATTGACTG CCCCCACTCA GCGTTTATAG ACGGAATATT TTATTTGTAT GTGTTCTAGA CTTTCTTGCT
1901 CTTTAGAAAA CTTCCTTTCA TTCAGCATCT CTGGCAGATA CACACTCCCA TTTTCTCCTT TTCAGATCAA CTACACAATT CTCTCCTGTT GCTTATGTTC
2001 CCCTGTGCCT GGGCGAAGCT TCTACTCCAC GTAGGCACCA AGACGGATGG ACGTCAGTGC TCACATCTCT CATCCAGACG AATAGTTGCT GTTCCATAAT
2101 TTAATTTCAC AGCAGCATTG GACACAAACA GCCAGAAATT TGGCAGTGAT GACCTTCACT GTTCATCTTC ACAATTTATG ATGAGGGATC AAATAAATTT
2201 TATTAAGGTT TATTCATCCT GTTTAAGAGA TTGTCCCTGA AAAAGGCTTG CTTGTAAAAA AATAACTAAA ATTCTGTTAC AGTTTGAACC AAGATGCCAC
2301 GAAGCAAATC CTTTGTATTT TTCCATTTTT AATGCTTTTG TTACTGATGT GTTGGCATGA GTTCATAATC ACGCTGATCT AGTAAAATGA CCTTCGTTAG
2401 AATGGTCCTT TTCTTGGGAC TCAGAGACCC CTTCATTCTC CTGAGCTTCC CCCATCCCAC TTTTGCTCTT CCCATAACTT AACAAAGTTT CCCATGCCCT
2501 TGAGTTGAGT GTATTTGTTT GTTTGTTTGT TTTATTAACA TCATCTGAGT ACCACCATTC TTTTCCACAG CCTGTGCAAG CAGGTGGATG TTGGGACTTC
2601 ACATATCAGG CAACTTTAAC CTACAAGAAA TCACGTGTGT TCTCAAGATC CTGTCCAGAA AGAAAGAAAG AGAGAAAGAG AGAAAGAGAG AAAGAGAGAG
2701 AGAGAGAGAG AGAGAGAGAG AGAGAGAGAG AGAGAGAAAG AGAGAGAGAG AGAAAGAAAG CAGCAACAAT ATAATTTTGA GTTATGACCC AGGACATCAA
2801 TAGTTGAACA TGGCGATGAC TCACAGAAGT AATGGCATTC AAAGTTAGGA TGCAGACTAC ACTCCTAGAA CAAACAGACC AAGAACAGCA TACTTGGATA
2901 GCTCACAAGC CCTTTCTAAA ACTATGCATT CTCTCAGCCC AGCCCTCTGT CCTTGTTCCT GTCTGTCTGT CTGTTTCTCA TCTTACTCAG TTAAAGGCAG
3001 CACAATCCAC TTACTGACCA GCTGGGGAAT TAAAACAAAC ACTTCTTTTA TTTCCTATCG TCTTCTCTTT CAAAGATACT GCCACTCTGC AACCTCCTTC
3101 CAATTCTATC GTGAACATCC GGCTACTGCC CACACGCCAT CCACTCATTC TCTTGTCTTC TGGATTATCT CCTTTGCCAC CCCTCCATCG TATTTCAGTA
3201 AGATTGATCT CTTCAAGGAT CAAATGTGAG CTTTCGTATT GCCAAGCTCT TCTCTGATTT AAAGCCAACT CCCTTTACCA CCCTCCTAAG TAGTTAGGCA
3301 TTCCCTTCAG GACAAAGAAA TTTCATTTCA TTTCATTAAG AGCATCGTCG CTACCCTTGA ATATGTATGT CGTTTTAAAG CTTAGAAGGC CCCTTGCCCA
3401 TCTCCTTCCT CTGCCAAGTC ACCACTCAGA ATTTATGGCT TACTTTAGAT TTCTGCTTCC TTCCACGTTG TTTCCAGCCC ACATACAGTG TGTCTCTGGC
3501 ATCTGCCCTC AGGAGACTTC AACTTCTCTA ATGATATCGG TTGGCTACTC AAATGCAAAA TCTGTCCTAT CCCCTGCCCC GTATCTTGGA ACTTTTATGG
3601 GTTCTGACTT CTACTTATTT GACATTCTAT CGCTAGGTCC TAGCCAAGGA GGACCCATGA ATGGTTAGAA ATAACCAAGG AGCTGCAGCA ACGAGAGGAT
3701 GGATGCATGA ATTAATGAAA TTGCCTCCAG GGCCTGGAGA CTAGAGAGGG GAGTGTTGGG CAAAAGTGTT TGTGTAGTGA ATGGCAAAGT CAGCTGGGGT
3801 GTTGGTTAGG CACAGTAAAA AGGGGTTATT TGACACTTGT TACTTTTCAC CTTCTTAAAA CAGAAATGAG TATATGTGAA GGTTAAGGTC AATGTTATGT
3901 TATACTGGAA TTTAGAAATT CACTGTGAAG GTTACACTTA CTGACCAGCA AGCCCAAGAC CCTCCTGTCT CCTGTCCTCC AGAACCCAGA TTCTAGCTAA
4001 GCATTAGGAA ACCTAATCAT TTCTCCTGGC CCCGACACTT CATTTTCAGT CAATCTAATG TATAATAATA CCTCTAAAGT CAATACCTAC TTCCTAGTTT
4101 GGTTCAAAAT GCTTTCCATA GAAGGGTGCT GATCATGGGC CTCACGTTAA AGACAGGCCT TATGTGGCCC AGAGCTTGCT GGCACACTCC GTTGGAGTGA
4201 ACTATTCAAG CCTCCCAGGG CTGCTTAAGC CACAGCCTCC TTGCCAACTA TCCAGACAGA AGCTCAGAGC ATCCCTAAGA GGCTGTGGGG GAGGAGAAGC
4301 CGATTATCAA GTTTCATGTA CAAAAGAGAA AAAGTGAGCA GTAACTGGAA AGTCCTTAAT CCTTCTGGGA ATTCTGGCTA ACGGGTCAGG AATTAACCTT
                                                TATA_m                          ↓m
4401 GGCGGTAATT AGGCCATCAG TGACTCTTTC AGAGAATGTT CAAAAACCTC AGCTTTGTTT GGAGAACTCG TGTGGGTGGG AGCAGCTGTT TGCACAAAGA
     TATA_M                    ↓M
4501 ACAGCATAAA GTCTAGCTGC TCCTCACCAC GTCCAGGACC AGATCTCTGG CGGGGAGTAG GGTGCAAGGC AGCCAGGTCT CCGATCCTTT CCGCAGCATG
4601 GGGAAGCACA GGCTCCGGAG CCTGGCCCTG CTGCCACTGC TGCTGCGGCT GCTGCTGCTG CTGCTGCCCA CCGATGCCTC AGCTCCACAA AAACCGTGAG  Ex 1
4701 TCCAGCAAAC GCCAGACGCT TCGATTTAGG GTTTCTGTTT ACTCGTTTCT TAAAACCATG ACCTTCTTTC TCATCTCCGA GATTCTGAGA TTTTTCCAAA
4801 CCTTCCTCAG TCATTCTTCC TCCTCAATAC AGTTTTAATC ATTTCTCTCT AGCTCTCTGG ATTTCGCGAT ATCCAAACTC CCCCACTGGA CATGTAGGAC
4901 CTTTCTAGGT ACTTGTGTCT TCAGATCCTA ATGTGGCTCT GTTGAGACTT TGTCCTTGCT GCCCCCTTTA CATTCCTCTG TCGGTTTGAC AGAAAAGACA
5001 CGCCTCTCCA GCGTGCGTTT TGTCCCTGGG TTTAAAATCC CCTAGGAATG ATTTCAAACG TTGATGTTGA TGCTGAAAAT GCAAACGACT TTTGTTATGG
5101 AGAAAAACTA TTGTGTGCAA ATAAGCTAGC GATGGATTCT GTGTTTCAAA GAAATGAACC AAGGAGAGTA GACAGCCAGA CAGATGGCTG GAACAATACT
5201 TGGGAGCACT GGTTCTGGCT GATTTAGAGC AGGTTATACA AGAGGAACGG TGACAAAGAG GTAGTTACTG GAGGCTTTTT CTTCGGGCCT ATATTTACAG
```

```
5301  ATAAGACGAA TGTTATTGTT TTAATGAGGC AAACCAGCTC AGGGTCATCA TCAATGCTCA CGGTGGGAGA GGAGCACACA CAATTCTCAA CAGAGATGTG
5401  CTGTGAGAGT ATCTCACTTT TACGTCTAAT ATATATGGGC CATGAACATA TTTTTAGTAA GTTATGTACT AGAACTTCTG AAAATAATCC AATCATAAGA
5501  AAGCATTTAA CACTTACTAA TTTATGGGCA CATATATAAA ACAGGAGTTA CTGTTAGACT TACGTATCCA CAGGTGATAT TTCTTGTACA TTTATCTGTT
5601  AATAAAAAAA AATGAGAATT GTTCATGAGT TATTTTGCAG CATAAAGAAA TCATAGTATA AACTCTGGAT AATAGTAAGT GGCACAAAAA TTAGAATTTT
5701  TTGAGCATTA CAAAAAGAAA CCTATACATT TTATTTATTT ATTATTTTCT CTCGTTGTTC ATTTGTTTGT TTTGAGACAG GATTTTGCTA AGGGTGCCTG
5801  ACAGCATCGA ACTTGTGGAA ATCCTCCTGT CTCAGCCTCT CAAGTCATGA GCCACATCTC TCAATATTTA TTGTTAAGCA GGTGATACTA CCTATAATAT
5901  TTAATATAAT AATTCTGTGG AGAAGTTGAA ATCTAACTTA AGAGAAATTA AGCCTTATAT TAGGTTATCA AAAGACTTTC CTGTTTGAAA TTGTAATAAA
6001  AATAGAAGTT AACTGGAGTG TTACATAAGA GTCCTATGGC ATTTAAGAAC ATTCTTGAGA GTTGCTGTTT GAAACTGGAG TGTAAGTCTT TATCATACAT
6101  TATTGATAAG CTGTCTGCCC TGTTCCTGAG TTACCAGAGG ATGGATGACG TTTGAATGGA GCTTTTCTCT ACATTTGGAA AGGCAAATCC ACTTGATGCA
6201  AGCATCCAGG TTGGGTAGCC CACACAGACT TAGTTAACTG CAACTGCAGC GAGACTCTCC ACTAGTGACT CTGTTTTCGT GAGGCTCAGC TTCCTTAGTT
6301  ATAAAATAGA GGCAGGTTCT TGATGTGGAA GTTTTACAGG CTTGCTGAAT ACATGTGGGT GCCTTCTTAT CCCCCACCAC CCTCTCAACA AAGACGTTCA
6401  CTGCAATGGC CTGAGGGGGG GAGGGGATTC TTAGGCCATC CACAGCCGAG TCTGCCTTCT CCTACTCAAC AGCTAAGGAT GATGATGGTG GCACCGTGAT
6501  GGGACTCAGA GATGCTCCTA GCAACAACCC CTTCCCTCTC CTCTAACTGC ATTCTTCCAC CCTCAGAATC TACATGGTGA TGGTTCCCTC CCTGCTCCAC
6601  GCAGGAACCC CCGAGAAGGC CTGCTTCCTG TTCAGCCATC TAAACGAGAC AGTGGCTGTG AGAGTGTCCT TGGAGTCTGT CCATGGGAAC CAAAGCCTCT  Ex 2
6701  TCACTGACCT TGTAGTTGAC AAGGACCTAT TCCACTGTAC CTCCTTCACC GTGAGTCAAG GGGTGGGAGC TTCGGGATGG TAGGCGTGTC AATAAAGAAT
6801  GTAACTATGG CAACCAAGTG TAGGGATTAA GGATCAACTC ATCTGTTTCC AACTGCATAT TAGAAGCATT TTGAGAATAA ATCCAATGAC CAGAGAAAAA
6901  ATTCTATTTG TCAAAACAAC AACAATGATG ACGACAATGA GGATGACAGA AACAGCAGGC AGTAGAGAAG GGAAGAAATG TGACTTTTAA AATTAAAATT
7001  CCTTTTTAAA GTTCATAAAG AAAGCTAGTT TAAAGGTAGC CATAAGCCAC CACTGGTAGT TGACAGGCTA CCTCAGACTC TTTTAAGAGT AAAAATTGTA
7101  GAAATCAACC TGAGCAAGAG TCCAATATAT CTTCAGAGCT GGATGTAACA GGGAATACTT GTAATCTAGT ACTGCCAAAT GCTTTCTTTT TCCACAGTGT
7201  TAGACCTGAT TTCACAGTGT TCTGTTAGTG TATAATTACC TGGTCAAATA CAGTGTGTAA CCGAGTCAAG ATTGTACTGT GTCTTCTATA CAAACCGAGA
7301  TGACTTTTTC AGTTGTAGCT CTTCTTAGTA ATGTATTTCC TCACCTTAAA CTTGGGTCTT CCAATTTGTA AAATATAAGA ATCCTTGCCT ATGCCAACTT
7401  CCCAGAAGAG AATGAAGGAG TGAAATCAAG AAAAAGAATG AGATCTCCAT AGAGATATAG AACAGGGGGA GGAGAAAGAG GCTGGAGAGA CTTTCCACGT
7501  TTTCAGACTC TCTTCTGTGT TTTAGGTCCC ACAGTCTTCA TCTGATGAGC TGATGTTTTT CACTGTCCAA GTAAAAGGAG CAACTCATGA GTTCAGGAGG
7601  CAGAGCACGG TGCTGGTTAA GAAGAAAGAG AGCCTGGTCT TTGCTCAGAC TGACAAGCCC ATCTACAAAC CAGGACAGAC AGGTATGATG AGGTCCACAA  Ex 3
7701  TCAGGATAGC AGCAAAAAGG AAACTCCTCT CTCCTTGCCT GTCCCTCTCA GCCCCTGGGA TCC
```

FIGURE 4: DNA sequence of the 5′-terminal part of the α_2M gene. The map positions of the 5′ ends of the minor and major RNA size classes are indicated by arrows labeled m and M, and the corresponding TATA boxes are underlined and labeled $TATA_m$ and $TATA_M$. Exons 1, 2, and 3 are underlined and identified by the labels Ex1, Ex2, and Ex3 in the margin.

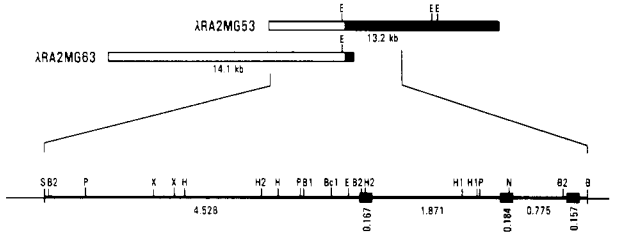

FIGURE 5: Block structure and restriction map of the 5′-terminal region of the α_2M gene. Boxes in the two top rows represent the inserts of genomic clones λRA2MG-53 and -63; the insert sizes in kb are given below the boxes. Dark areas are transcribed sequences. An enlarged segment corresponding to the sequence given in Figure 4 is shown in the bottom row. Exons are symbolized by black boxes. The sizes of the exons (in kb) and of the introns are given underneath. Restriction enzyme cleavage sites are abbreviated as follows: E, *Eco*RI; B1, *Bgl*1; B2, *Bgl*2; Bc1, *Bcl*I; B, *Bam*HI; H, *Hin*dIII; H2, *Hpa*2; HI, *Hpa*I; N, *Nco*I; P, *Pst*I; S, *Sal*I; X, *Xba*I.

oligonucleotide F representing a sequence located between the major and minor transcription start sites. With oligonucleotide B, two RNaseH-resistant 5′ RNA fragments were found in the livers of rats during an acute-phase response. The lengths of these resistant fragments [m(358) and M(295), Figure 9, lane 3] permitted the evaluation of the locations of the 5′ ends of the corresponding RNA species, and these locations coincided within the resolution limits of this technique (±5 nucleotides) with the positions shown in Figure 6A. With oligonucleotide F, only transcripts originating from the minor transcription start site were detected (Figure 7, lane 6). This experiment provides independent evidence that both the minor and major transcription start sites are used in rat liver, that the concentration of RNA molecules of both size classes is greatly induced under acute-phase conditions, and that the majority of transcripts originate from the major start site. Again, with this technique no α_2M RNA was detected in livers from control animals (Figure 7, lanes 1 and 4), indicating that this RNA species is of very low abundance in the livers of normal rats.

Dual transcription start sites are utilized by a number of other genes, often in a mutually exclusive manner, to achieve cell type specific expression in hepatocytes and other tissues (Schibler et al., 1983; Perlino et al., 1987; Schibler & Sierra, 1987). We have tested the hypothesis that liver hepatocytes might preferentially utilize the major transcription start site and that liver macrophages, the Kupffer cells, which are known to express the α_2M gene, might preferentially utilize the minor start site. Therefore, polyadenylated RNA was prepared from the rat hepatoma cell line H35, after treatment with glucocorticoids and monokines under conditions known to stimulate the synthesis of α_2M RNA in these cells (Baumann et al., 1987). Then RNaseH experiments were performed as described above to analyze the transcripts from the major and minor start sites. We found that both transcription start sites are used in the clonally derived hepatoma cell line H35 (Figure 8, track 7). Therefore, hepatocytes utilized both start sites, and thus the hypothesis of cell-type specific differential promoter utilization was wrong. The same result was obtained when RNA from primary rat hepatocyte cultures was used, which were enriched for hepatocytes. The expression of α_2M in these cultures was stimulated by treatment with glucocorticoids and macrophage supernatants containing IL6/HSF (data not shown).

In pregnant rats the α_2M gene is expressed in the uterus and the placenta (Hayashida et al., 1986). We have investigated whether both transcription start sites are also used in

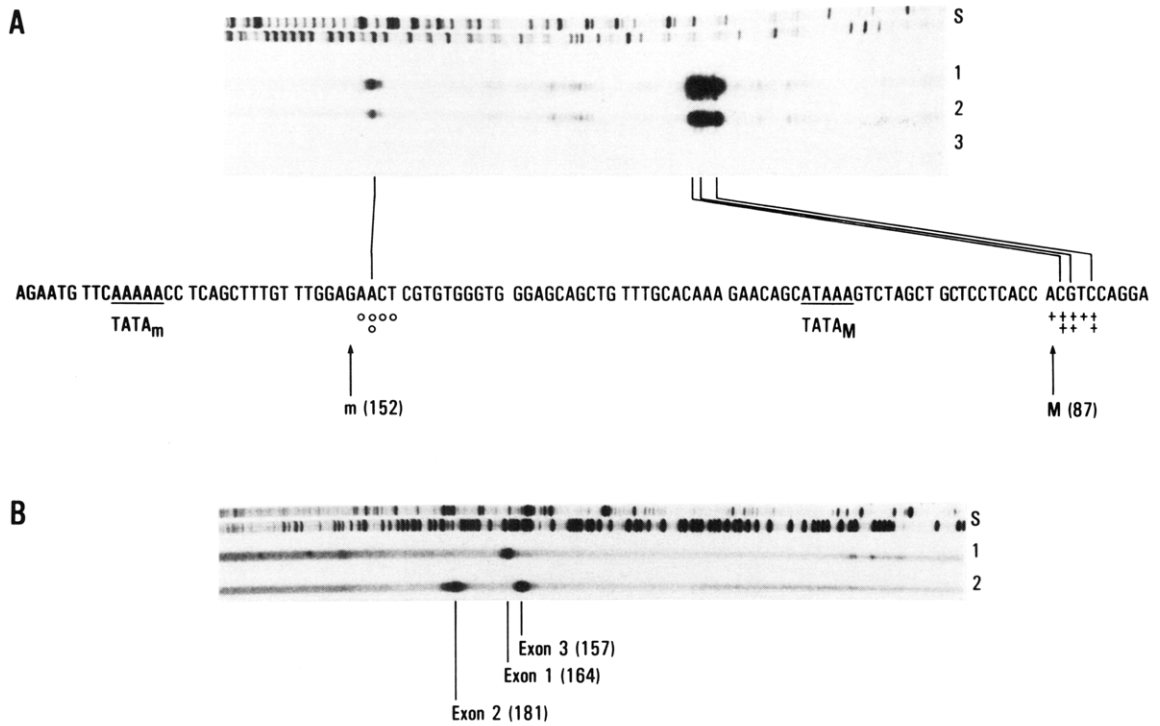

FIGURE 6: Definition of the 5′ ends of α_2M RNA by S1 mapping. (A) S1 mapping using single end-labeled genomic DNA fragments. A 1.515-kb *Hin*dIII/*Sal*I fragment of clone λRA2MG-63, containing exon 1 (the *Sal*I site is artificial, originating from the *Sal*I linker at the end of the insert) was isolated and digested with the enzyme *Hpa*II resulting in a cut at a unique site of exon 1. The 5′ ends were labeled with polynucleotide kinase and [γ-^{32}P]ATP. The fragment was subcut asymmetrically with the enzyme *Pst*I at a unique site, and the large 928-bp *Pst*I/*Hpa*II fragment was isolated, carrying a single end label at its *Hpa*II end in exon 1. This fragment was denatured and annealed with size-enriched, polyadenylated RNA from control rats and rats with an experimentally induced inflammation. S1 mapping was performed as described under Experimental Procedures. Tracks S, size marker (T and C tracks of a known DNA sequence); tracks 1 and 2, 6 and 3 μg, respectively, of RNA from 18-h inflamed rats; track 3, 5 μg of control rat liver RNA. The strong and weak bands on the autoradiograph are labeled by ‡ and +, respectively, and by ° and o in the sequence. Arrows point to the deduced positions of the 5′ ends of the major [M(87)] and the minor [m(152)] RNA size classes. The numbers 87 and 152 refer to the distance in nucleotides from the labeled end of the DNA fragment. TATA$_M$ and TATA$_m$ designate the positions of the corresponding TATA boxes. (B) S1 mapping using internally labeled genomic DNA. The 1.555-kb *Pst*I/*Bam*HI fragment of clone λRA2MG-53 containing exons 2 and 3 and a 0.56-kb *Eco*RI/*Sal*I fragment of clone λRA2MG-63 containing exon 1 and part of the 5′-flanking sequence were subcloned into the vector M13mp8, and internally labeled single-stranded DNA, complementary to the mRNA sequence, was prepared as described (Northemann et al., 1988b). The single-stranded DNA was used for S1 mapping under the conditions described under Experimental Procedures. Tracks S: size marker; track 1, genomic DNA containing the exon 1 sequence [the length of exon 1 (164 bp) was determined by comparison with the size marker]; track 2, genomic DNA containing exons 2 and 3 [the deduced lengths of exons 2 and 3 (181 and 157 bp, respectively) are given in parentheses].

the uterus and have prepared polyadenylated RNA from the uterus and the liver of the same pregnant females. The α_2M gene was strongly expressed in the uterus, at approximately 15–20 times greater abundance than in an acute-phase liver, and both RNA 5′ ends were used at comparable relative intensities as in the acute-phase liver (Figure 8, tracks 1–6). Thus, among the cell types tested, the major and minor transcription start sites were not utilized differentially.

No α_2M RNA was detected in the liver of pregnant females, suggesting different mechanisms of induction of the α_2M gene in acute-phase liver and in the uterus during pregnancy.

Discussion

The objective of this study was to isolate the rat α_2M gene, to characterize its 5′-terminal region, and to locate its transcription start site. These investigations are prerequisites for a more detailed analysis of the cis- and trans-acting control elements responsible for the tissue-specific expression and inflammatory regulation of this gene. The finding of two transcription start sites led to an investigation of their possible cell type specific utilization in different cell types within the liver, and in other α_2M-producing cell types in rats.

Only one type of restriction fragment pattern characteristic of the α_2M gene was found in the rat genome. No fragments were seen in the genomic DNA digest that were not represented in the clones. Theoretically, two or more gene loci could be present in the rat genome, but they would need to have indistinguishable *Eco*RI patterns, which is unlikely in view of the large size of this gene (48 kb). Therefore, we conclude, that rats have only one α_2M locus in analogy with the single α_2M locus present in the human genome (Kan et al., 1985) and in agreement with similar findings by others (Hayashida et al., 1986). This situation is different from the rat α_1I3 genes, which form a small gene family. Under similar experimental conditions as those used here, at least four different α_1I3 genes were found in the rat genome (Braciak et al., 1988; Northemann et al., 1988b). It is not known whether an equivalent of the α_1I3 gene family exists in the human genome, but from the sequence relatedness it is clear that rat α_2M is the equivalent of human α_2M. The human α_2M gene is not an acute-phase gene and is expressed at constitutively high levels. Once the control region of the human α_2M gene is known, it will be interesting to compare it with that of the rat α_2M gene in order to understand the acute-phase regulation of the rat gene in terms of its unique control sequences.

The transcription start sites of the α_2M gene have been mapped in this study with three independent techniques, producing results in good agreement within the error limits of the techniques. In addition to the primer extension experiments we have performed primer extension/RNA sequencing experiments and analyzed the reaction products on a sequencing gel (T. Braciak, W. Northemann, and G. Fey,

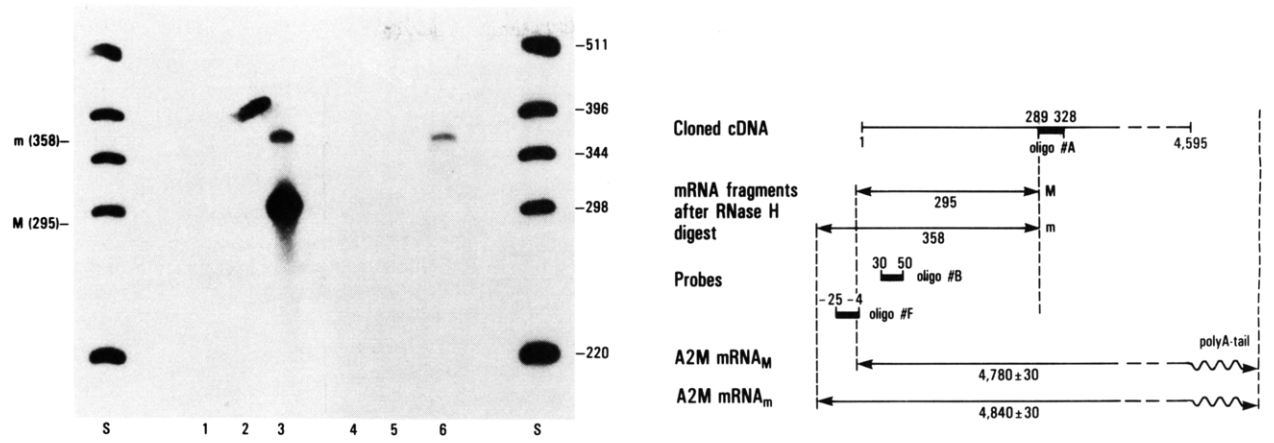

FIGURE 7: Both the minor and major RNA 5′ ends are utilized in liver cells; characterization of the 5′ ends by RNaseH mapping. Ten micrograms of size-enriched polyadenylated RNA from livers of control rats (tracks 1 and 4) and 6-h (tracks 2 and 5) and 18-h inflamed rats (tracks 3 and 6) was annealed with oligonucleotide A, corresponding to nucleotides 289–328 in the published cDNA (Gehring et al., 1987, see interpretative scheme on the right). The RNA/DNA heteroduplexes were treated with RNaseH as described under Experimental Procedures, and the RNaseH-resistant RNA fragments were electrophoretically separated in a 6% polyacrylamide–6 M urea sequencing gel. The RNA was transferred from the gel to a nylon membrane by electroblotting with a Bio-Rad trans blot cell. The filter was cut in half, and the left half(tracks 1–3) was hybridized with radiolabeled oligonucleotide B, representing cDNA sequence from nucleotides 30–50 (see scheme on the right). The right half (tracks 4–6) was hybridized with radiolabeled oligonucleotide F, corresponding to genomic DNA sequences located upstream of the cloned cDNA sequences (indicated by −4 to −25 in the scheme on the right). These are noncoding sequences from the first exon of the gene located between the map positions of the 5′ ends of the minor and major RNA size classes (see Figure 6A). RNaseH-resistant 5′ fragments originating from both RNA size classes were revealed with oligonucleotide B [labeled m(358) and M(295) in the left margin] (track 3), while only the fragment originating from the minor size class m(358) was revealed with oligonucleotide F (track 6). The numbers 358 and 295 refer to the length of those RNaseH-resistant fragments, which define the position of the 5′ ends of the minor and major RNA size classes, as shown in the scheme on the right.

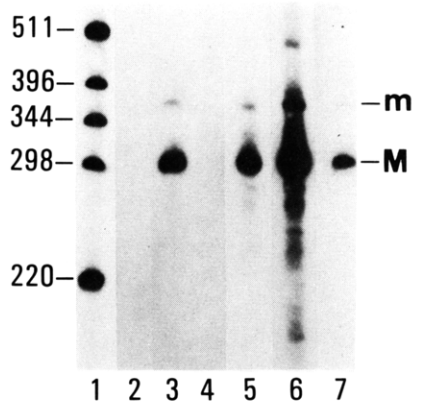

FIGURE 8: Equal utilization of both α_2M RNA 5′ ends in acute-phase livers, pregnant rat uterus, and clonally derived hepatoma cells. Size-enriched polyadenylated RNA was prepared from normal rat livers and 18-h inflamed rat livers. In addition, polyadenylated but not size-enriched RNA was prepared from the liver and the uterus of the same pregnant female and from H35 hepatoma cells (Baumann et al., 1987), treated with 2×10^{-6} M dexamethasone and 1000 units/mL recombinant human IL6 (Brakenhoff et al., 1987). The RNA was annealed with oligonucleotide A, treated with RNaseH, and analyzed by hybridization with radiolabeled oligonucleotide B as described in Figure 7 and Experimental Procedures. (Track 1) Size markers, sizes (in nucleotides) given on the left. (Track 2) 4 μg of size-enriched normal liver RNA. (Track 3) 4 μg of size-enriched acute-phase liver RNA. (Track 4) 15 μg of liver RNA of pregnant female. (Track 5) 15 μg of RNA from the uterus of a pregnant female, 4-h exposures of the autoradiograph. (Track 6) same as track 5 but 48-h exposure. All other tracks are 48-h exposure. From the differences in exposure time required to reach bands of equal intensity and the different amounts of RNA loaded, it was estimated that the α_2M RNA in pregnant uterus is 15–20 times more abundant than in acute-phase liver. (Track 7) 15 μg of polyadenylated RNA from H35 hepatoma cells. m and M in the right margin designate the RNaseH-resistant RNA fragments corresponding to the minor and major RNA size classes.

unpublished data). In this case the 5′ end was mapped at the CG pair one and two nucleotides downstream of the site labeled M(87) in Figure 6B. Therefore, our best data indicate that the 5′ end of the major RNA size class is located only 6 nucleotides upstream of the end of the published cloned cDNA sequence, and not 21 ± 5 nucleotides upstream, as previously reported (Gehring et al., 1987). The improvement is due to more precise size evaluation of the RNaseH-resistant RNA. The major α_2M mRNA species has a size of 4780 ● 30 nucleotides, including the poly(A) tail, and the less abundant species would be 4840 ± 30 nucleotides long (Figure 7).

It is likely that the 5′ ends of the stable α_2M mRNA species as they were mapped here represent indeed the transcriptional start sites of the gene. Formally the alternative cannot be ruled out, that transcription originated even further to the 5′ side and that the stable mRNA species resulted from the original transcripts by posttranscriptional trimming from the 5′ end. However, this is unlikely, because no well-characterized precedent for this mechanism has been reported in mammalian cells. In addition, we have accumulated further evidence by transfection of cloned α_2M DNA sequences into cultured hepatoma cells and by transient expression assays that functional promoters reside within ±50 bp of the major site mapped here (Hattori, Abraham, Bradshaw, and Fey, unpublished data). Finally, the transcription start regions of the α_2M gene and the closely related α_1I3 prototype gene (Northemann et al., 1988b) show a significant degree of sequence conservation: the start site of the α_1I3 gene coincides within ±1 nucleotide with the minor start site of the α_2M gene (Figure 9). Since both genes have evolved from a common ancestor, this observation indicates that the minor start site of the α_2M gene is a weak but genuine promoter.

Substantial concentrations of the α_2M protein circulate in the blood of pregnant rats. It has therefore been discussed that pregnancy might resemble an acute-phase response and that the circulating protein might be synthesized in the liver. Very large concentrations of α_2M mRNA (approximately 15–20 times higher than in the livers of male rats during an acute-phase response) were found in the uterus of pregnant females (Figure 8, lanes 5 and 6). Hayashida and colleagues (Hayashida et al., 1986) have previously reported similar

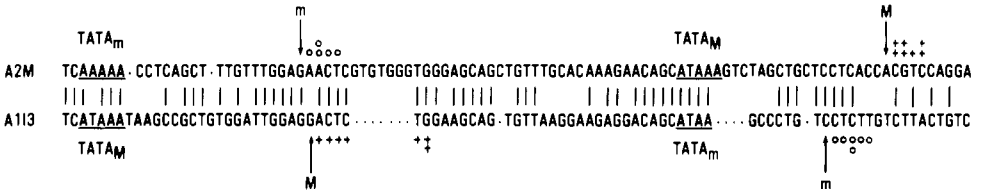

FIGURE 9: Alignment of the 5′-terminal regions of the α_2M and α_1I3 genes. The 5′-terminal regions of the α_2M gene (Figure 4) and the prototype α_1I13 gene (Braciak et al., 1988; Northemann et al., 1988b) were manually aligned. Gaps were introduced to achieve optimal alignment and are designated by dots. The map positions of the 5′ ends of the minor and major α_2M RNA classes are indicated by arrows, and the symbols are as in Figure 6A. The transcription start site of the α_1I3 gene is designated M. A potential minor start site of the α_1I3 gene is designated m (Northemann et al., 1988b).

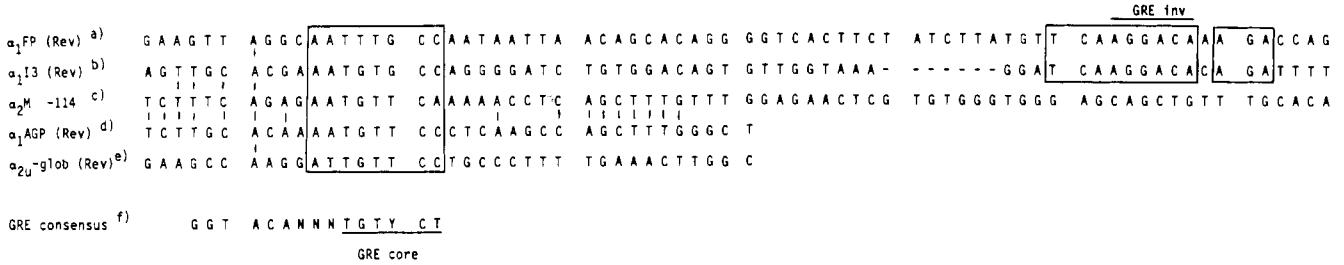

FIGURE 10: Location of a potential glucocorticoid responsive element and conserved sequences with the α_1-acid glycoprotein gene in the 5′-flanking sequence of the α_2M gene. The 5′-flanking region of the α_2M gene and the inverted sequences (Rev) of the α_1-fetoprotein gene (α_1FP, row 1), the α_1I3 gene (row 2), the α_1-acid glycoprotein gene (α_1AGP, row 4), and the α_{2u}-globulin gene (α_{2u}-glob, row 5) were aligned by using the UWGCG program BESTFIT (Devereux et al., 1984). A region strongly conserved between these genes is boxed in the left part of the figure. This area shows imperfect homology (5/6) with the GRE core hexanucleotide, TGTYCT, as defined by Jantzen et al. (1987). This region was shown to be functionally relevant sequence, which mediates the transcriptional regulation of the α_1AGP and α_{2u}-globulin genes by glucocorticoids (Baumann & Maquat, 1986; Addison & Kurtz, 1986). Additional conserved nucleotides between the α_2M and α_1AGP genes in this area are identified by vertical traits. Another GRE element, shared between the α_1FP and α_1I3 genes, is shown in the top right part of the Figure. Footnotes: (a) From Chevrette et al. (1987); (b) from Northemann et al. (1988b); (c) from Figure 4; (d) from Reinke and Feigelson (1985) and Baumann and Maquat (1986); (e) from Addison and Kurtz (1986); (f) GRE consensus sequence from Jantzen et al. (1987) (hexanucleotide core is underlined). Gaps in the α_1I3 sequence were introduced for optimal alignment (Northemann et al., 1988b).

results as well as high levels of α_2M mRNA in the placenta of pregnant females. To our surprise, no α_2M mRNA was found in the liver of pregnant females. Therefore, the α_2M protein circulating during pregnancy probably originates from the uterus or the placenta or yet other extrahepatic sites of synthesis, but not from the liver.

Since glucocorticoids participate significantly in the regulation of the α_2M gene during an acute-phase response (Northemann et al., 1988a; Baumann et al., 1986) and since consensus cis-acting control sequences mediating the effects of glucocorticoids are known, we have searched for such control elements in the 5′ region of the α_2M gene by computer-assisted DNA sequence inspection. Several different consensus sequences for the glucocorticoid responsive element (GRE) have been published (Karin et al., 1984; Moore et al., 1985; Jantzen et al., 1987). These conserved sequences of approximately 16 bp in length represent extended binding sites for the glucocorticoid receptor. All three published GRE consensus sequences coincide in a conserved hexanucleotide core, 5′-TGTYCT-3′, and differ slightly in the surrounding sequences. No perfect copy of an extended GRE consensus sequence was found in the 7.7 kb of 5′-terminal sequences of the α_2M gene. However, an imperfect copy of the hexanucleotide GRE core, conserved in 5/6 nucleotides, is present in the promoter region, at −97 to −102 nucleotides upstream of the major transcription start site (Figure 10). In this area, the sequences of the α_2M gene and the α_1-acid glycoprotein gene (α_1AGP; Reinke & Feigelson, 1985) are highly conserved. This region is known to mediate the glucocorticoid response of the α_1AGP gene (Baumann & Maquat, 1986), and this is therefore a plausible candidate for an element mediating the glucocorticoid control of the α_2M gene. The functional relevance of this conserved region needs to be established by functional assays. With the 5′-terminal sequence of the α_2M gene in hand, we can now begin to identify the functionally relevant sequences for the control of this gene by mediators of inflammation including IL6 and by other hormonal signals.

ACKNOWLEDGMENTS

We thank Drs. T. Sargent and G. Scherrer for kindly providing the rat genomic DNA libraries; Dr. K. Lonberg-Holm for encouragement and critical reading of the manuscript; Drs. G. Fuller and L. Abraham for gifts of RNA from primary hepatocytes and hepatoma cell lines; Dr. H. Baumann for the H35 cell line; G. Hudson for the synthesis of oligonucleotides; Drs. H. Pannekoek, L. Aarden, and J. Brakenhoff for generous gifts of purified recombinant human IL6; and Dr. M. Hattori for critical discussions. The help of H. Short, D. Kunz, and M. Heisig with the isolation and characterization of genomic DNA, of G. Jensen with the VAX computer, and of Keith Dunn for the preparation of the manuscript is gratefully acknowledged.

REFERENCES

Addison, W. R., & Kurtz, D. T. (1986) *Mol. Cell. Biol. 6*, 2334–2346.

Bankier, A. T., & Barrell, B. G. (1983) in *Techniques in Nucleic Acid Biochemistry* (Flavell, R. A., Ed.) Vol. B5, pp 1–34, Elsevier/North-Holland, Limerick, Ireland.

Baumann, H., & Maquat, L. E. (1986) *Mol. Cell. Biol. 6*, 2251–2561.

Baumann, H., Hill, R. E., Sauder, D. N., & Jahreis, G. P. (1986) *J. Cell Biol. 102*, 370–383.

Baumann, H., Onorato, V., Gauldie, J., & Jahreis, G. (1987) *J. Biol. Chem. 262*, 9756–9768.

Birch, H., & Schreiber, G. (1986) *J. Biol. Chem. 261*, 8077–8080.

Braciak, T. A., Northemann, W., Hudson, G. O., Shiels, B. R., Gehring, M. R., & Fey, G. H. (1988) *J. Biol. Chem. 263*, 3999–4012.

Brakenhoff, J. P. J., de Groot, E. R., Evers, R. F., Pannekoek, H., & Aarden, L. A. (1987) *J. Immunol. 139*, 4116–4121.

Chevrette, M., Guertin, M., Turcotte, B., & Belanger, L. (1987) *Nucleic Acids Res. 15*, 1338–1399.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C., & Chambon, P. (1980) *Science (Washington, D.C.) 209*, 1406–1414.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Esnard, F., & Gauthier, F. (1980) *Biochim. Biophys. Acta 614*, 553–563.

Fey, G. H., & Fuller, G. H. (1987) *Mol. Biol. Med. 4*, 323–338.

Gauldie, J., Richards, C., Harnish, D., Lansdorp, P., & Baumann, H. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7251–7255.

Gauthier, F., & Ohlsson, K. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 981–992.

Gehring, M. R., Shiels, B. R., Northemann, W., de Bruijn, M. H. L., Kan, C.-C., Chain, A. C., Noonan, D. J., & Fey, G. H. (1987) *J. Biol. Chem. 262*, 446–454.

Gordon, A. H. (1976) *Biochem. J. 159*, 643–650.

Gordon, A. H., & Koj, A. (1985) *Res. Monogr. Cell Tissue Physiol. 10*.

Hayashida, K., Okubo, H., Noguchi, M., Yoshida, H., Kanagawa, K., Matsuo, H., & Sakaki, Y. (1985) *J. Biol. Chem. 260*, 14224–14229.

Hayashida, K., Tsuchiya, Y., Kurokawa, S., Hattori, M., Ishibashi, H., Okubo, H., & Sakaki, Y. (1986) *J. Biochem. (Tokyo) 100*, 989–993.

Jantzen, H.-M., Strähle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R., & Schütz, G. (1987) *Cell (Cambridge, Mass.) 49*, 19–38.

Kan, C.-C., Solomon, E., Belt, K. T., Chain, A. C., Hiorns, L. R., & Fey, G. H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2282–2286.

Karin, M., Haslinger, A., Holtgreve, A., Richards, R. I., Krauter, P., Westphal, H. M., & Beato, M. (1984) *Nature (London) 308*, 513–519.

Lonberg-Holm, K., Reed, D. L., Hebert, R. C., Roberts, R. C., Hillman, M. C., & Kutney, R. M. (1986) *J. Biol. Chem. 261*, 438–445.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY.

Moore, D. D., Marks, A. R., Buckley, D. I., Kapler, G., Payvar, F., & Goodman, H. M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 699–702.

Northemann, W., Heisig, M., Kunz, D., & Heinrich, P. C. (1985) *J. Biol. Chem. 260*, 6200–6205.

Northemann, W., Ueberberg, H., & Fey, G. (1988a) *Life Sci. Adv. 7*, 7–10.

Northemann, W., Shiels, B. R., Braciak, T. A., & Fey, G. H. (1988b) *Biochemistry* (in press).

Perlino, E., Cortese, R., & Ciliberto, G. (1987) *EMBO J. 6*, 2767–2771.

Reinke, R., & Feigelson, P. (1985) *J. Biol. Chem. 260*, 4397–4403.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5476.

Sargent, T. D., Wu, J. R., Sala-Trepat, J. M., Wallace, R. B., Reyes, A. A., & Bonner, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3256–3260.

Sargent, T. D., Jagodzinski, L. L., Yang, M., & Bonner, J. (1981) *Mol. Cell. Biol. 1*, 871–883.

Schauefele, J. T., & Koo, P. H. (1982) *Biochem. Biophys. Res. Commun. 108*, 1–7.

Schibler, U., & Sierra, F. (1987) *Annu. Rev. Genet. 21*, 237–258.

Schibler, U., Hagenbüchle, O., Wellauer, P., & Pittet, A.-C. (1983) *Cell (Cambridge, Mass.) 33*, 501–508.

Schmid, W., Scherer, G., Danesch, U., Zantgraf, H., Matthias, P., Strange, C. M., Röwekamp, W., & Schütz, G. (1982) *EMBO J. 1*, 1287–1293.

Schreiber, G. (1987) in *The Plasma Proteins* (Putnam, F. W., Ed.) Vol. 5, Academic, Orlando, FL.

Shiels, B. R., Northemann, W., Gehring, M. R., & Fey, G. H. (1987) *J. Biol. Chem. 262*, 12826–12831.

Soberon, X., Covarrubias, L., & Bolivar, F. (1980) *Gene 9*, 287–305.

Sottrup-Jensen, L. (1987) in *The Plasma Proteins* (Putnam, F. W., Ed.) Vol. 5, pp 191–291, Academic, Orlando, FL.

Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Lønblad, P. B., Jones, C. M., Wierzbicki, D. M., Magnusson, S., Domdey, H., Wetzel, R. A., Lundwall, Å., Tack, B. F., & Fey, G. H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 9–13.

Staden, R. (1986) *Nucleic Acids Res. 14*, 217–232.

Turpen, T. H., & Griffith, O. M. (1986) *BioTechniques 4*, 11–15.