Sandhoff, K., Schwarzmann, G., Sarmientos, F., & Conzelmann, E. (1987) in *Gangliosides and Modulation of Neuronal Functions* (Rahmann, H., Ed.) NATO ASI Series Vol. H7, pp 231-250, Springer-Verlag, Berlin.
Schwarzmann, G., Hoffmann-Bleihauer, P., Shubert, J., Sandhoff, K., & Marsh, D. (1983) *Biochemistry 22*, 5041-5048.
Sonderfeld, S., Conzelmann, E., Schwarzmann, G., Burg, J., Hinrichs, U., & Sandhoff, K. (1985) *Eur. J. Biochem. 149*, 247-255.
Sonnino, S., Ghidoni, R., Marchesini, S., & Tettamanti, G. (1979) *J. Neurochem. 33*, 117-121.
Sonnino, S., Ghidoni, R., Fiorilli, A., Venerando, B., & Tettamanti, G. (1984) *J. Neurosci. Res. 12*, 193-204.
Sonnino, S., Kirshner, G., Ghidoni, R., Acquotti, D., & Tettamanti, G. (1985) *J. Lipid Res. 26*, 248-257.
Sonnino, S., Acquotti, D., Riboni, L., Giuliani, A., Kirschner, G., & Tettamanti, G. (1986) *Chem. Phys. Lipids 42*, 3-26.
Spiegel, S., Schlessinger, J., & Fishman, P. H. (1984) *J. Cell Biol. 99*, 699-704.
Staab, H. A. (1962) *Angew. Chem. 74*, 407-410.
Staros, J. V. (1980) *Trends Biochem. Sci. (Pers. Ed.) 5*, 320-322.
Svennerholm, L. (1957) *Biochim. Biophys. Acta 13*, 856-858.
Svennerholm, L. (1970) in *Handbook of Neurochemistry* (Lajta, A., Ed.) pp 425-452, Plenum, New York.
Tettamanti, G., Bonali, F., Marchesini, S., & Zambotti, V. (1973) *Biochim. Biophys. Acta 296*, 160-170.
Tomasi, M., Roda, G., Ausiello, C., D'Agnolo, G., Venerando, B., Ghidoni, R., Sonnino, S., & Tettamanti, G. (1980) *Eur. J. Biochem. 111*, 315-324.
Venerando, B., Roberti, S., Sonnino, S., Fiorilli, A., & Tettamanti, G. (1982) *Biochim. Biophys. Acta 692*, 18-26.
Wiegandt, H. (1985) *New Compr. Biochem. 10*, 199-260.

Structure and Negative Transcriptional Regulation by Glucocorticoids of the Acute-Phase Rat α_1-Inhibitor III Gene[†,‡]

Wolfgang Northemann, Brian R. Shiels,[§] Todd A. Braciak, and Georg H. Fey*
Department of Immunology, Research Institute of Scripps Clinic, La Jolla, California 92037
Received March 24, 1988; Revised Manuscript Received August 5, 1988

ABSTRACT: DNA clones representing the negative acute-phase gene coding for the plasma proteinase inhibitor α_1-inhibitor III were isolated from a rat genomic library. Structural analysis established the existence of at least four different members of the α_1-inhibitor III gene family. Partial DNA sequence analysis of the 5′-terminal regions was performed for the α_1-inhibitor III gene and the related α_1-inhibitor IV gene. The transcription start site of the α_1-inhibitor III gene was located by S1 mapping and primer extension. No stable α_1-inhibitor IV mRNA was detected in rat liver. In an experimentally induced acute-phase reaction, the transcription rate of the α_1-inhibitor III gene was reduced 12.7-fold at 6 h after stimulation. Four hours after injection of a high dose of dexamethasone into rats, the transcription rate of this gene was reduced 9-fold. Thus, glucocorticoids alone are capable of causing a strong transient down-regulation of the transcription of this gene in rats, independent of other inflammatory mediators. An inverted consensus glucocorticoid responsive element (5′GG$^{A}_{C}$A$^{C}_{A}$A$^{T}_{A}$T3′) shared with the glucocorticoid-regulated α_1-fetoprotein, α_{2u}-globulin, and α_1-acid glycoprotein genes was detected by computer-assisted sequence analysis in the promoter proximal 5′-flanking region of the α_1-inhibitor III gene.

With normal concentrations of 6-10 mg/mL, α_1-inhibitor III (α_1I3)[1] is the second most abundant plasma globulin in rats after albumin (Gauthier & Ohlsson, 1978; Esnard & Gauthier, 1980; Esnard et al., 1985; Lonberg-Holm et al., 1987). It is a thiol ester protein, and together with the related α_1-macroglobulin (α_1M) and α_2-macroglobulin (α_2M), it is part of a protein family that also includes complement components C3, C4 and C5 (Sottrup-Jensen et al., 1985; Sottrup-Jensen, 1987; Gehring et al., 1987; Braciak et al., 1988). While α_2M is the most dramatically increased acute-phase protein in rats, α_1M and C3 are increased only about 2-fold during an acute-phase response, and α_1I3 is among the most strongly down-regulated acute-phase proteins in rats. During the first few days of an acute-phase response, α_1I3 concentrations fall to 1-2 mg/mL, and during chronic inflammations, they further decline to less than 0.5 mg/mL (Lonberg-Holm et al., 1987).

We have recently isolated and sequenced α_1I3 cDNA clones and deduced the α_1I3 protein sequence (Braciak et al., 1988). The mature α_1I3 polypeptide is 1453 amino acids in length

[†] This work was supported by Grants AI22166 and AI19651 from the National Institute of Allergy and Infectious Diseases and by an award from the United Liver Association (to G.H.F.). Short-term fellowships from the North Atlantic Treaty Organization (to W.N.) and the Fulbright Commission (to B.R.S.) provided additional support. This is Publication No. 5056-IMM from the Department of Immunology, Research Institute of Scripps Clinic.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J03552.
* Correspondence should be addressed to this author at the Department of Immunology, IMM14, Research Institute of Scripps Clinic, 10666 North Torrey Pines Rd., La Jolla, CA 92037.
[§] Present address: Welcome Unit of Molecular Parasitology, Department of Veterinary Parasitology, Bearsden Rd., Glasgow G61 1QH Scotland, United Kingdom.

[1] Abbreviations: α_1I3, α_1-inhibitor III; α_1I4, α_1-inhibitor IV; α_2M, α_2-macroglobulin; α_1M, α_1-macroglobulin; α_1AGP, α_1-acid glycoprotein; α_1FP, α_1-fetoprotein; SDS, sodium dodecyl sulfate; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); bp, base pair(s); kbp, kilobase pair(s); HSF, hepatocyte stimulating factor; IL-6, interleukin 6; GRE, glucocorticoid receptor element.

0006-2960/89/0428-0084$01.50/0

with a calculated molecular weight of 161 053. In that study, we also reported the observation of at least four structurally distinct variant classes of α_1I3 liver mRNA species. This observation was unexpected, because heterogeneity of α_1I3 had not been previously detected at the plasma protein level. We concluded that the most likely explanation was the existence of a small α_1I3 gene family. Here we provide definitive proof for this hypothesis through the isolation and structural characterization of five different types of α_1I3 genomic DNA clones. It is not yet known whether all of these clones represent functional genes. We were unable to detect cytoplasmic mRNA in rat livers for one member of this family, called α_1-inhibitor IV (α_1I4). However, several members of this family are simultaneously expressed in rat liver, because corresponding liver mRNA species were detected in the form of cloned cDNA both by our group and others (Braciak et al., 1988; Schweizer et al., 1987; Aiello et al., 1988; Feigelson, personal communication).

α_1I3 is a plasma proteinase inhibitor (Gauthier & Ohlsson, 1978; Esnard & Gauthier, 1980; Sottrup-Jensen, 1987). From the variant cDNA sequences, we deduced that different members of the α-macroglobulin family and the α_1I3 subfamily possess different bait regions and predicted they may inhibit different spectra of proteinases (Gehring et al., 1987; Braciak et al., 1988). α-Macroglobulins are multifunctional polypeptides. Apart from serving as proteinase inhibitors, they also function as carrier proteins for metal ions, small basic polypeptides, and hormones such as platelet-derived growth factor (Barrett, 1981; Huang et al., 1984). An additional function for α_1I3 as a general clearance protein has recently been suggested (Braciak et al., 1988). Human equivalents of α_1I3 are unknown, but the nucleotide sequences of rat α_1I3 presented here may become useful tools in the search for the human homologues and their functions.

We have isolated rat α_1I3 genomic DNA clones to study the mechanism of the transcriptional down-regulation of this gene during the acute-phase response. The α-macroglobulin gene family is an excellent model for such studies because the closely related members of this family are regulated very differently by inflammatory mediators. The comparison of the control mechanisms of the different members of this family is expected to facilitate the identification of the cis- and trans-acting acute-phase responsive control elements of these genes. In a parallel study, we have isolated the rat α_2M gene and initiated the characterization of its transcriptional control region (Northemann et al., 1988b). The transcription increase of the α_2M gene occurring during an acute-phase response is dependent on the action of glucocorticoids and other inflammatory mediators (Northemann et al., 1988a), while other acute-phase genes, such as the rat α_1-acid glycoprotein gene, can be regulated independently by both classes of agents (Kulkarni et al., 1985a,b; Baumann & Maquat, 1986; Klein et al., 1987). Therefore, we have asked whether the effects of both groups of agents on the expression of the α_1I3 gene can be separated. Here we have investigated the effect of glucocorticoids alone on the transcription of this gene in living rats and have compared them with the expression changes that occur during an acute inflammatory response.

Experimental Procedures

Animals and Materials. Male Fisher 344 rats were obtained from Simonsen Laboratories, Gilroy, CA. Primer extension/RNA sequencing was performed with the Gem Seq transcription sequence system from Promega. Synthetic oligonucleotides were produced on an Applied Biosystems Model 380A DNA synthesizer. Prior to their use as primers, oligonucleotides were separated electrophoretically in polyacrylamide–urea gels and gel-eluted.

Isolation and Mapping of Genomic DNA Clones. A total of 500 000 plaques were plated for each of the two rat genomic DNA libraries, the *Eco*RI partial digest library (Sargent et al., 1979) and the *Hae*III partial digest library (Sargent et al., 1981). Plaque lifts were performed on Genescreen nylon membranes. The filters were hybridized in a first round with nick-translated cDNA fragments representing all of the sequences coding for the mature α_1I3 prototype polypeptide (gel-eluted *Bam*HI restriction fragments of cDNA clone pRLA1I3/23J; Braciak et al., 1988). Hybridization was performed in 50% formamide, 20 mM piperazine-*N*,*N*′-bis-(2-ethanesulfonic acid) (PIPES), 0.8 M sodium chloride, 1 mM EDTA, and 0.1% sodium dodecyl sulfate (SDS), containing 100 μg/mL single-stranded salmon sperm DNA at pH 6.5, 42 °C, for 24 h. The filters were washed and exposed for autoradiography by using standard procedures (Maniatis et al., 1982). The first-round isolates were rescreened with a nick-translated cDNA probe representing a 585 bp *Bam*HI fragment from the 5′ end of the insert of clone pRLA1I3/23J (Braciak et al., 1988). λ phages resulting from this second screen were purified to homogeneity, and restriction maps for the enzymes *Eco*RI, *Bam*HI, *Kpn*I, *Xba*I, and *Hin*dIII were prepared for the phages λRA1I3G-3, γRA1I3G-8, and λRA1I3G-21 by standard procedures (Maniatis et al., 1982).

Southern Blot Analysis. Preparation of DNA from adult rat liver, digestion with *Eco*RI, electrophoresis, and Southern blot analysis were performed by standard procedures (Maniatis et al., 1982). The DNA was transferred to a Genescreen membrane by electroblotting in 25 mM sodium phosphate, pH 6.5, using a Bio-Rad transblot cell and fixed on the membrane by illumination for 3 min with ultraviolet light as described (Gehring et al., 1987). The membrane was prehybridized and hybridized with 200 ng of a kinase-labeled oligonucleotide probe (cDNA nucleotides 208–241; Braciak et al., 1988), washed, and exposed for autoradiography under standard conditions (Maniatis et al., 1982).

DNA Sequence Analysis. Three adjacent *Eco*RI fragments of 2.68, 1.27, and 1.23 kbp of clone λRA1I3G-3 comprising the 5′-terminal region of the α_1I3 gene and two adjacent *Xba*I fragments of 2.04 and 2.06 kbp of clone λRA1I3G-21 representing the corresponding region of the α_1I4 gene were subcloned in M13 phage vectors. Double-stranded replicative intermediate DNA was prepared; the inserts were recovered in microgram quantities and randomly subcloned into M13 vectors by established procedures (Bankier & Barrell, 1983). DNA sequencing was performed by using the dideoxy sequencing technique (Sanger et al., 1977; Bankier & Barrell, 1983). DNA sequence data were collected, aligned, and analyzed using Staden's DB system and ANALYSEQ programs (Staden, 1986) and the University of Wisconsin Genetics Computer Group programs (Devereux et al., 1984) on a Digital VAX 11/750 computer. Sequencing of random clones was continued until each character of the final sequence was covered at least once, preferably twice on each strand. On average, each nucleotide was read from six to seven independently sequenced M13 clones, resulting in databases of 53 043 and 19 869 characters of primary sequence data for the 5.18 kbp α_1I3 gene segment and the 4.09 kbp α_1I4 gene segment, respectively.

Experimental inflammation and mRNA extraction as well as size selection of polyadenylated mRNA by sucrose gradient sedimentation were performed as previously described (Gehring et al., 1987; Braciak et al., 1988).

S1 Mapping. A 1.1 kbp *Kpn*I/*Eco*RI genomic DNA fragment containing exon 1 of the α_1I3 gene (Figure 4) and two fragments each containing exon 1 of the α_1I4 gene (a 2.06 kbp *Xba*I/*Xba*I fragment and a 0.59 kbp *Nco*I/*Sau*3A fragment) were subcloned into the phage vector M13 mp8. The constructs were verified by partial DNA sequencing. Single white plaques (infected colonies) were isolated for those constructs that produced single-stranded phage with inserts complementary to the α_1I3 mRNA. *Escherichia coli* JM101-TG1 cells infected with these phages were grown in liquid cultures as described (Bankier & Barrell, 1983) in the presence of 1 mCi of [^{32}P]orthophosphate. Internally labeled, single-stranded phage DNA was prepared according to standard procedures (Bankier & Barrell, 1983). One-half microgram of ^{32}P-labeled DNA was annealed with 10 μg of size-enriched, polyadenylated mRNA over 4000 nucleotides in length. The mRNA was prepared from livers both of normal rats and of rats 18 h after intraperitoneal injection of complete Freund's adjuvant (CFA). The annealing reaction was performed in a 20-μL volume in 60% formamide, 10 mM PIPES, pH 6.4, 0.4 M sodium chloride, and 1 mM EDTA for 14 h at 50 °C. To stop the reaction, the annealing mix was chilled on ice and diluted with 180 μL of ice-cold S1 buffer containing 280 mM sodium chloride, 4.5 mM zinc acetate, and 30 mM sodium acetate, pH 4.4. The sample was then treated with 600 units of S1 nuclease in the presence of 5 μg of single-stranded salmon sperm DNA for 45 min at 37 °C. The nucleic acids were then extracted with phenol and precipitated with ethanol. The precipitate was redissolved in 7 μL of 10 mM Tris/0.1 mM EDTA, pH 8.0, and a 3-μL aliquot was applied to a sequencing gel (Bankier & Barrell, 1983).

Primer Extension/RNA Sequencing. The 30-nucleotide signal peptide oligonucleotide underlined in Figure 3A was used as a primer. Ten nanograms of the primer was annealed for 10 min at 60 °C and for an additional 30 min at 42 °C with size-enriched mRNA (over 3000 nucleotides in length) from rats 18 h after induction of an experimental inflammation, according to protocols provided with the Gemseq transcript sequencing system (Promega). Then 2.5 μL (25 μCi) of [α-^{35}S]thio-dATP was added to the annealing mixture, and the extension and stop reactions were performed following the instructions of the supplier of the Gemseq kit (Promega Notes, number 7, January 1987). Aliquots (3 μL) were electrophoresed in sequencing gels.

Transcription Rate Measurements and Dot Blot Analysis of Nuclear RNA Species. Nuclear run-on experiments were performed as published (Northemann et al., 1985). The target DNA sequences immobilized on the filters were the plasmids pRSA13 and pRSA57, each carrying approximately 1200 bp of cDNA sequence inserts representing the 3′ and 5′ parts of the rat albumin mRNA sequence (Sargent et al., 1979). The plasmid DNA was linearized by partial digestion with *Pst*I. The α_2-macroglobulin plasmid pRLA2M/29 (Gehring et al., 1987) and the α_1I3 plasmid pRLA1I3/23J (Braciak et al., 1988) contained approximately 4600 bp of cDNA sequence inserts each. These plasmids were linearized by digestion with *Bam*HI. Fifteen micrograms of plasmid DNA was immobilized on each filter. RNA was extracted from isolated nuclei with guanidinium thiocyanate and purified by centrifugation through a 5.7 M cesium chloride cushion as described (Shiels et al., 1987). The purified RNA was resuspended in sterile water and stored at −70 °C. For dot hybridization, 15 μg of RNA (per dot) was diluted in 50 μL of 25 mM sodium phosphate buffer, pH 5.5, and heat denatured for 15 min at 65 °C. The samples were then chilled on ice, and 100 μL of ice-cold 25 mM sodium phosphate buffer, pH 5.5, was added. The RNA was deposited on a Genescreen nylon membrane using a Schleicher & Schuell minifold apparatus. The nucleic acids were fixed on the membrane by illumination with ultraviolet light as described (Gehring et al., 1987). Hybridization with radiolabeled probes specific for the first intron of the α_1I3 and the third intron of the α_1AGP gene and wash conditions were as published (Shiels et al., 1987).

Results

Rat Genome Contains a Family of α_1I3-Related Genes. An unexpectedly large number of 148 genomic DNA clones were isolated by screening only 2 complexities of each of 2 different rat genomic DNA libraries with cDNA probes representing the coding sequences for the α_1I3 protein (Braciak et al., 1988; Experimental Procedures). To focus on the 5′ end of the gene, these isolates were rehybridized with a cDNA fragment representing the 5′-proximal part of the α_1I3 mRNA. Of the 11 resulting clones, 9 had different restriction cleavage patterns, and 3 were identical. The nine clones fell into five groups with different *Eco*RI cleavage patterns (Figure 1C, groups I–V). Group I (clones λRA1I3G-3, -8, and -46) carries a characteristic 2.7 kbp *Eco*RI fragment that hybridized with an oligonucleotide probe from the signal peptide coding sequence of the α_1I3 prototype (Braciak et al., 1988; Figure 1, panel B, tracks 1, 2, and 7). Group II (clone λRA1I3G-16) differs from group I by a long *Eco*RI fragment at the 3′ end of its insert. Group III (λRA1I3G-21) shows a 10.5 kbp *Eco*RI fragment hybridizing with the signal peptide probe (Figure 1B, track 4). We have designated the gene represented by this clone as the α_1-inhibitor IV (α_1I4) gene. Group IV clones λRA1I3G-38 and λRA1I3G-45 carry a 5.8 kbp *Eco*RI fragment hybridizing with the signal peptide probe, which is truncated in clone λRA1I3G-38 (Figure 1B, tracks 5 and 6). Finally, group V includes clones λRA1I3G-53 and λRA1I3G-54, with a similarly sized 5.8 kbp *Eco*RI fragment as group IV, but differing from group IV clones in the size of another *Eco*RI fragment, located at the 3′ end of the insert in clone λRA1I3G-53 (Figure 1B, tracks 8 and 9, Figure 1C). From the restriction patterns, the probe sequences, and the hybridization conditions used, we deduced that none of these isolates were either α_2M or α_1M clones.

Southern blot analysis of rat genomic DNA digested with the restriction enzyme *Eco*RI was performed under stringent conditions of hybridization (Figure 2) and revealed four fragments of 10.5, 5.8, 3.7, and 1.27 kbp hybridizing with an α_1I3 gene exon 2 specific 34-mer oligonucleotide probe (nucleotides 208–241 in the α_1I3 prototype cDNA sequence; Braciak et al., 1988). Two of these were the characteristic *Eco*RI fragments of groups III and IV/V, respectively, as described above. Under the conditions of hybridization used, this probe does not cross-hybridize with α_2M sequences. The α_1I3 exon 2 specific probe also differs substantially in its sequence from the known α_1M cDNA sequence (Eggertsen, Shiels, Hudson, and Fey, unpublished data) and is thus unlikely to cross-hybridize with α_1M sequences. Therefore, we conclude that the genomic fragments revealed in Figure 2, track 1, are α_1I3 specific. Thus, the rat genome contains an α_1I3 gene family with at least four different members positioned at a minimum of two different gene loci.

DNA Sequence of the 5′-Terminal Regions of the α_1I3 and α_1I4 Genes. The DNA sequences of a 5.18 kbp section from the 5′-terminal region of the group I genomic clone λRA1I3G-3 and of a corresponding 4.09 kbp section from the group III (α_1I4 gene) clone λRA1I3G-21 were determined by random subcloning into M13 vectors and dideoxynucleotide

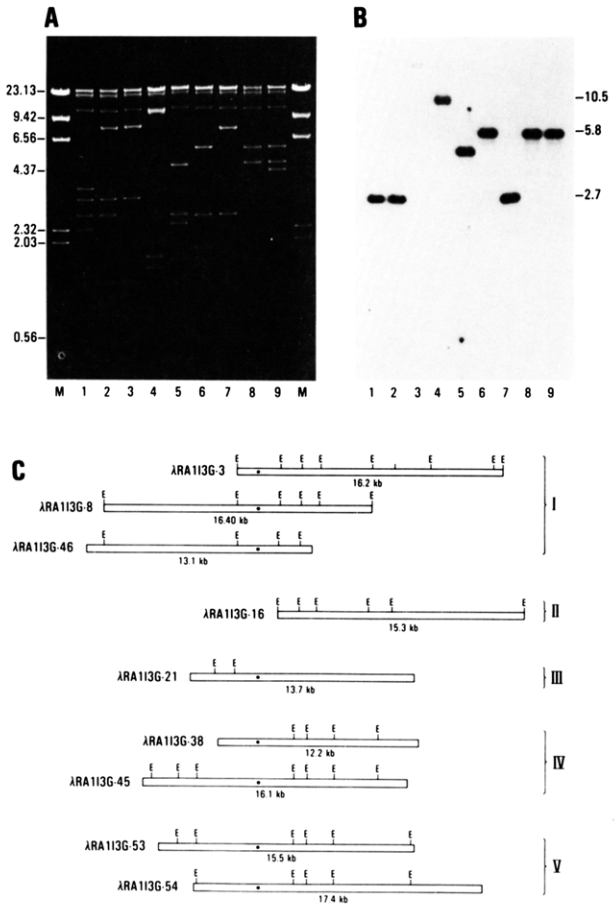

FIGURE 1: Restriction fragment mapping of α_1I3 genomic DNA clones. (A) Ethidium bromide stained genomic DNA fragments, separated in a 1% agarose gel. M: size marker, phage λ DNA digested with *Hin*dIII, sizes of the marker fragments (in kbp) are indicated on the left side. Tracks 1–9: Restriction fragments from clones λRA1I3G-3, -8, -16, -21, -38, -45, -46, -53, and -54, generated with the enzyme *Eco*RI. (B) Autoradiography of a Southern blot from a similar gel of *Eco*RI fragments as shown in panel A. Tracks 1–9: Same order of clones as in panel A. The gel was treated with 0.25 M HCl to facilitate the transfer of high molecular weight fragments. DNA was denatured and reneutralized in the gel by standard procedures (Maniatis et al., 1982) and transferred to a genescreen membrane by capillary blotting in 25 mM sodium phosphate buffer, pH 6.5, as prescribed by the supplier (NEN/Dupont, Genescreen users manual). The blot was hybridized with a 30-mer synthetic oligonucleotide representing signal peptide coding sequences (underlined in exon 1, Figure 3A), which was 5′ end labeled with [γ-^{32}P]ATP by polynucleotide kinase. Clone λRA1I3G-16 hybridized with the cDNA search probe in the original screening procedure, but not with this signal peptide oligonucleotide. Sizes of the characteristic hybridizing bands (in kbp) are indicated on the right. (C) Restriction maps for the enzyme *Eco*RI (E) of the five different groups of α_1I3-related genomic DNA clones. The inserts contained in all nine isolates are shown as open boxes. The total insert lengths in kilobases (kb) are given under each insert. Roman numbers on the right refer to the five groups of different restriction patterns as described in the text. Asterisks identify the location of the signal peptide oligonucleotide used as hybridization probe in panel B. Clone λRA1I3G-21 has a 10.5 kbp *Eco*RI inserted fragment bounded by one natural *Eco*RI site and one linker site. The genomic DNA carries a natural *Eco*RI site in close proximity downstream of this linker site, which gives rise to the natural 10.5 kbp *Eco*RI fragment visible in panel B, track 5.

sequencing (Figure 3A,B). The 5.18 kbp region contains the first three exons, the first two introns, part of intron 3, and 2.2 kbp of 5′-flanking sequence of the α_1I3 gene. The exon sequences of this gene were identical with the published prototype α_1I3 cDNA sequence of cDNA clone pRLA1I3/2J (Braciak et al., 1988), and thus group I genomic clones represent the prototype α_1I3 cDNA sequence. The 4.09 kbp

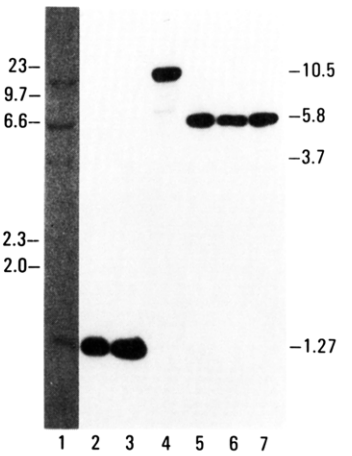

FIGURE 2: Southern blot analysis of rat genomic DNA. Rat genomic DNA (track 1) and purified DNA from the genomic phages λRA1I3G-3, -8, -21, -45, -53, and -54 (tracks 2–7) were digested with the enzyme *Eco*RI, electrophoresed in an agarose gel, and transferred to a genescreen membrane. The blot was hybridized with a 34-mer oligonucleotide probe specific for α_1I3 (cDNA nucleotides 208–241; Braciak et al., 1988) that was labeled with [γ^{32}P]ATP and polynucleotide kinase. Sizes of λDNA marker fragments produced by digestion with *Hin*dIII (in kbp) are on the left; sizes of characteristic *Eco*RI fragments of different members of the α_1I3 gene family (in kbp) are on the right.

Table I: Sequence Similarity between the 5′-Terminal Regions of the α_1I3 and α_1I4 Genes

region compared	% identity	region compared	% identity
5′ FS[a]	90.3	exon 2	98.9
exon 1[a]	90.0	intron 2	92.5
intron 1	90.4	exon 3	98.1

[a] The lengths of the 5′-flanking sequence (FS) and of exon 1 of the α_1I4 gene are defined by analogy with the α_1I3 gene. They were not determined directly, because the α_1I4 gene is not expressed in liver. Sequence similarities were calculated by the University of Wisconsin program BESTFIT (Devereux et al., 1984).

region of the α_1I4 gene contains exons 2 and 3, introns 1 and 2, part of intron 3, and an area homologous in sequence to exon 1 of the α_1I3 prototype gene. However, the 5′ boundary of exon 1 of the α_1I4 gene has not been determined, because no stable mRNA products from this gene were found in liver (see below). In addition, this fragment included sequences equivalent to approximately 1 kbp of the 5′-flanking region of the α_1I3 gene. The exon–intron block structure of these two fragments is schematically summarized in Figure 4. Both genes are closely related in their sequence (Table I). Exons 2 and 3 show 98.9% and 98.1% identity, respectively, while introns 1 and 2 are 90.4% and 92.5% identical. The first 1000 nucleotides of the α_1I3 5′-flanking sequence are 90.3% conserved in the corresponding region of the α_1I4 gene. The lengths of exons 2 and 3 identical, and the lengths of introns 1 and 2 are within ±20 bp of each other. Therefore, the members of the α_1I3 gene family share a much greater degree of sequence identity with each other (over 90% in the exon sequences) than with the coding sequences of other α-macroglobulin genes (α_1I3 and α_2M protein sequences are 58% identical; Braciak et al., 1988; Gehring et al., 1987).

Mapping the Transcription Start Site of the α_1I3 Prototype Gene. The transcription start site of the α_1I3 prototype gene (group I clone λRA1I3G-3) was mapped by a combination of S1 mapping (Figure 5) and primer extension/RNA sequencing techniques (Figure 6). The length of exon 1 of the α_1I3 prototype gene was found to be 169 ● 4 nucleotides (Figure 5, tracks 3 and 6). mRNA derived from this gene was present in normal adult liver (Figure 5, track 3) and in

A

```
   1  AATTCATTAC GACTTAAGAG CCTTGCATGG CAATGGTGTG CCTATAATCC TACTTCTATG GAGTTAGAAG GACCAAATCC CTTGGGATTC CTGTGTTAGA
 101  TGGTTTAACC TAATTTACAA CTTTCAGCAT CAAGAAAAAT GTCTCAATAA GCAAGGTAAA AATGACCCAA GACCTCTGTC CTCTTTACAA ACAAGTAAAC
 201  ACATGCAAGT GCTCATAAAA ACATAGACAT ATGTACTCAG ATACACTCAC ACAGACACAG ACACACAGAC ACACACAGAC ACAGACACAG ACACAGACAC
 301  ACACACACAC AGAGAGACAC ACAAGGCATG AAAGAGAGGA AAAAGAGAAA AACATTGATT AACATTTAAT CTTTCTGAAA TGACAGTTTT TCCCAGCACC
 401  AGATACTCAT GATGCCAGGA CAGAATCTGA GAGCACTCTC AAGATGTAGC TATAAATTAA ATTAGTAAGG GGTTTATAAA GGCAAAACTA ACAAGGGTAT
 501  GTACTTAATG CTTAAGTGCA GCCAGCCACA CATCCGGCCT ACATGCAGAA AAGCACACAT CTTTCATGCT GTCAGAGGCA AGCAAACTTG TTTACAGGAG
 601  TGAAAACACG TGATGTGTTA TCTTACATAA ACAAATCAGG TAATTCCAGG TATAATTCTC ACATTCTAGG AATGTCATCT GTCCTAGGGC AAGTGGGGCT
 701  CACAGGTTAG AGCCATTTTG TTTTGTATTA ATTTTTAAAC TTTTAAACAT AATATCCACC ATTGCACAAT ATTCTATCTT TTCTGATAGT TTTTGTTTGT
 801  TTTCAGTATG TCTGGCAAAT ATGTGTCTCT TGATCGACAT ACCTTAGACA TAACTGATAT TTTTAAGTTA AAATGCAAAA CTGCCTCATG TTGTTTTTGT
 901  TTCTTGTATT AGGATCTTCA CACATAACTC GAAGCACTAA AATGTACAAT AAATCTTGAT GCATGTCCTG TTTCTCCCAC ACCTGGGTGT CTCTGACTCG
1001  CATGGTCAGT GTGAGAGACC AGACCAACTC CAGCCTGGGT TCAGGCTCCA TCTTAGAGAG CCTGGCCTAA GGTTGAACTC AAGTTACCCA GCAAGCCTGC
1101  ATCTAGCACT AGTTTCCTAT TGATTAATTT GGGTTCAGAG GTCCAGCAAT GAAATACATA AGTTTTATGT TGTTTTTCCT TTCAAACCCG GGCATATCTC
1201  TGTGAAAGCT CTAGCTGTCC TGAACTAACT TTGTAGAGCA GGCTCTCCTC ATGAACTCCC AGAGATGTGC CTGCCTCTGC CTCCTGAGTT TTGGGGCTAA
1301  AGGCATGCAC CAGCCTGCCA GGCATAAGTT TTATTTTTTT CATCCCTTAC ATGTTTCCAT TGCATGCCCT TCTAGCATGT TCTTTCTTTC TACTCAATGC
1401  CTCACCATTA ATGTTGATTC AGCTTATGCT ACGCCTCACT CTCAAGCCAA TGTCACAGGA TCTTTGCAAA ACTAAACAGC CCTATGACCT AGACTAGAAA
1501  TTCTAATGTG GTTCCCAACT CATACTTAAT AACCTTCCAA AGGAACCGAG CTGCGTCCCA AATACTGGTA CCTCATCATT TTACACAATC ACCTGTAGTT
1601  GCACAGGTTA CACTTGGCCT CCTAGTAAAC CGCAGAGATA GATGCATGCT TCTAGATGAG GTTTCTGTTC TAACAAAGAA GATTCATGCA CAGTGACTTG
1701  TAGCAGAGGA GTTTTTGAGC TTGCCCAGTT TCCACAGTCC CCCTTTTATA CACAAGACAC AGAAAAGTAG ATACTTGAAT GATCAGAGCC AGGAGAACTT
1801  AAAATCTCCA ACCTGGACCA CTCTAGTGAC TCTCAAATAA GAAAGGAGAG AAACGGAGAT GCCACATCCC AGAGCGTGCT ACTAGATTCT AAGTGAGTAG
1901  AAGAGAGTGA GGTGGTAAAT GTCACACAAT CATAGCAGAG TCATGCACGG AGAAGGCTTA CTGGTCATTA AATTGAGAAA GTCTGCTCTA ATTCAGACTA
2001  AGCAAAATCT GTGTCCTTGA TCCTTTACCA ACACTGTCCA CAGATCCCCT GGCACATTTC GTGCAACTGG TGTAATTTAT TTGAGTTGTG CTCAGGCCAG
2101  CACTGAACAC AGTGACATTC ACAGAGAACA AATATTCTAT TTTGCTATTT GAACACCTTG GGCAGAGTTC TGTTTACAGA GAGGACTCAT AAATAAGCCG
2201  CTGTGGATTG GAGGACTCTG GAAGCAGTGT TAAGGAAGAG GACAGCATAA GCCCTGTCCT CTTGTCTTAC TGTCGTCCTG TCCTCCTGTC CTTCCATCAT
2301  GAAGAAGGAC AGAGAGGCTC AGCTGTGCCT TTTTTCAGCT CTTCTTGCCT TCCTGCCTTT TGCTTCCTTG CTCAATGGAA ACTCGTGAGT ACCACCAAGA
2401  CCCTCTTCCA CATCATCCCT CCTCCAATCG GAAACTATCC ATTTCCATTC ACAGTTCCTT TCCCTGAAAA ACTCACAAAA TTCTTTTTAC TTTTTAATTA
2501  ACAATTATGT GTGTGTATAT GTATATATGA TTGCAATTAT ACACATACAA TTCTTACCTG ATTCTAGGTG ATCTTTTTTC CAATTCTAAA GCAACCAGTG
2601  TTACTAAACT GTGTGGTATG ATTTAGAATA GTATTCCTCT TATTACTGAA CAAGGAGTGT CTCAAATCCC AATCATCAGA GAATTCTCAG GTGTTGAAAC
2701  TATTTAAGCA CTTTGTAAAT AAGCATTTTA ATGGGTTGGA AATGTGTAGC CTATGGTTCA ACCATGAGAC GGATGAGATG TCTGGCACAT CAGAACCCTG
2801  TGGATGTTTG TGTCCTAAAA CAATACTCTT GTTACGATCA AAGACTTTGG ATTAAACTAC CAATAGAACT TTTTATTTGA TGGTCAAAAT AACCATATAA
2901  TTAGGATTAT ATTGTTGAGT GATTTTATTT GTATCTGTGA CCAGAAAAAT ATATTTAAGG TGGTAAAATT AAAAAGCAAA CTAGTGGTAT TTTTGCATAG
3001  TGTGATATAG CAATGACTGG CTAGATTTGG AGGGCAAACA TCCCTAATTA TAGGAAAAGA TAAAGAATGA TTCCCACAGA CAGACTTCGG CACTTGCTCA
3101  AGTGTAACAT AGGTAAATTA TGCAAGTTTG AATTTCACAG TCAGGTCATA AACCTTTATT TGAAAAGTAC CCCAAAATAA AAAATAGGGT ATTGAGTTAT
3201  TTTTAAGGAT TAAAATGGTA GTGAAAAACA ATCATGATAA ATTTCTATGA AAAGAACAAT TATTCCTCCT GGATTTTCAT TTAAAAAATA AACCTTTCTA
3301  GAACGTGCAT GCTGGTTGCT TAGTGACAGG AGTGCACGCC ATTCAAACCT GCCCTGGCAG GTATCACTGC TGAAGGTGGT GATTGACGAG CTAATAATTG
3401  TCTTGCCACT CTGCTGTGTG ACAAACTGTG ACACTGCCAC ACTGTCAACC TGACATCTCA GACTGTTGAA ACATGGATGG TAATGGTGAC TGTGATGAAA
3501  CTCCAGTTAT AGACAATGAC ATCTGTCGTC AGCTGTCCGG AACCCCTTCC TTTCCAGGAA GTATATGGTG CTGGTCCCGT CCCAGCTCTA CACCGAGACC
3601  CCTGAGAAAA TCTGCCTGCA TCTATACCAT CTGAATGAGA CAGTGACTGT GACAGCTTCC TTGATATCTC AAAGGGGAAC GAGAAAACTG TTCGATGAGC
3701  TTGTGGTTGA CAAGGACTTG TTCCACTGTG TTTCCTTCAC TGTGAGTACT AAGTATCCGT CCTGATTTGT TTACATACTT ATGCACATGC GTGTGAGTAT
3801  GTACATACAA TACCCATACA TGTTTCCAAA AAATGTTATT CTAGAAACAT ATCTGTGATC TCAAAACAAT AGGATTTTAG AACATTTACC CAATAAACAA
3901  ATAGAATACA TCATTCCACC ACAAAATGAG AAGTCAAAAC TTGTCCCATT GAATTCCCTC AAAAAGCCCA CAGACCAGAG AAAGAAAGGG TTCTCCAGTT
4001  CATAGCACTT CCTGCTAATA AGTTAATAAA CTTTCAAACA GCACTCCAGA TAATTAGGTT CACAGCTCCT ATGCATAATC CTGAGATCCT GAATAAGTGA
4101  TGGGTTTCTT CTTTGGGACC TATCCTCAAG ACATAATTCT TACCACCATT AATACATTAT CTTATTCTCT ACCTCAGTTT GAGAGTTAGA GCTAAGTTTC
4201  TTTTGTGTCT TAGAGGTAAT TTTCAATGAA GATTATATTT ACTCTGTAAA GATGACTGTG TTTTTGTTGT TCACATAAAC TCCCACTGAC ATACAGAAAA
4301  TGATTTTTCT TTTCATCCCT TTCTTTTCCT GGCAATTTTT CCCTCCATGT AATACTTAAA CAGAAACTGA TGGTGCCTCC ACCTCCATCA GAGAAAAGGT
4401  TAATGAGGTA CTGTAAGAAA GGCAGAGGGA GATGGTGAAG GGTCTGTGTT CCTGTCGTTC TCAGATCCCT AGATTGCCCT CTTCTGAAGA GGAGGAGTCA
4501  CTCGACATCA ACATTGAAGG GGCAAAGCAT AAATTCAGCG AAAGGCGTGT GGTGCTTGTG AAGAACAAAG AAAGTGTTGT CTTTGTCCAG ACTGATAAGC
4601  CCATGTACAA GCCAGGACAG TCAGGTATGA AGCTCCACGG TGGGCAAAGG AATTAATACA CCACCCTAAA CATGGTGTGA AGAAGGCCTG GGTCTTTGGA
4701  ACTCACTAGA AAATTCTCTT TTTCCAGTTA AATTTCGGGT TGTCTCTATG GATAAAAATC TACACCCCCT GAATGAGTTG GTGAGTTTTT TACCTATGGG
4801  TATTGGTGGG TTCTTCAAGA CAAAATGTCA CAGCAACATT CTAATTTCCA GCTCCTTCTG AGACGTGGGT TAGCCCTTTT TTCTATTCTG TTCCTTTGCC
4901  ACTTTATAGC CTCAGCTCTT CTACCATTGG CCATCCTATG GCTTCCTTAT ATGTGACTGT CCTTATTTGG ACTCCATTTG ACTGGATAAG TATCTCTGGA
5001  TTGTCTTCAT ACTCTGTGGA GCCTTATACT GTATAGAACC TTTTAGGTTT TTCAGTCATT TCTTTTCTAG ATTCTGTTAC CATCCCTCAG AAATCTATTC
5101  CTAGTGTGAC AGGGATATTT TTCACACAGT CTCTATTATT TTATAACCTC ATATTTAACA CTTGGGTTCT TCCCTAGATG AATT
```

Ex1

Ex2

Ex3

B

```
   1 CTAGATGTCC TGGACTGGCT TTGTAGACCA GGCTCTCCTC TGGAACTCAT ATAGATATTC CTGCTTCTGC CTCCTGAGTT TTGGGGCTAA AGGCGTGCAC
 101 CAGCCTGCCA GGCCCAACTT TTTTCATCCC TTACATGCTT CCATTACATT CCCTTCTAGC ATGTTCTTTC TTTCTACTCA GTGTCTCAAC ATTAAAGTTG
 201 ATTCAGCTTA TGCTATGCCT CCCTCTCAAG CCAGTGTCAT GAGATCTTTG CAAAAATAAA CAGCCCTGTG ACCTAGACTG GATATTCTAA TGTGGATTCC
 301 AACTCCTACC TTAATAACCT TCCAAAGGAA CCAAGCTTCG TCAGAAATAC TGAAACTTCA ACAATTGACA CAATCACCTG TAGTTGCACA GGTTATACTT
 401 GGCCTCCTAC TAAACCGCAG AGATAGAAGA TCAGGTTTGT TGACTTCTGT TCTAACAAAG ATTTAGGAAC AGTGACTTGT AACAGGGGAG TTTTTGAGCT
 501 TGCCCAGTAT CCACAGTCCC CTTTTTATAC ACAGGACCCA GAAAAGTAGA TACTTGAATG ACCAGAGCCA GTAGAACTTA AATCTCTCCA ACCTGGACCA
 601 CTGTAGTGAC TCTCAAATGA GAAAGGACAG AAATGGAGAT GACACATCCC AGAGCGTGCT ACTAGATTCT AAGTGAGTAG AAGAGAGTGA GGTGGCAAAT
 701 GTCACACAAT CATAGCAGAG CCACCCATGG AGAAGACTTA CTGGTCATTA AGAAGAGAAA GTCTGCTCTA ATTCAGACTA AGCAAATTCT GTGGCCTTGA
 801 CCCTTTACCA ACACTGTCCA CAGATTCCCT GGCACCTTTC CTGCGACTGT TGTAATTTAT TTGAGTTGTG CTCTGGCCAC CACTGAACAC AGTGACATTC
 901 ACAGAGAACA AACATTCTGT TTTGCTATTT GAACACCTTG GGCAGAGTGT TCTGTTTAGA GAGAGGACTT ATAAATAAGC CACTGTGGAT TGGAGGACTC
1001 TGGAAACGGT GTTAAGGAAG AGGATAGCAT AAGCCCTGTC CGCTTGTCTT CCTGTTGTCC TATCCTCCTG TCCTTCCACG ATGTGGAAGA ACAGAGAGGC  Ex1
1101 TCACCTGTGC CTTTTTTCAG TTCTTCTTGC CCTCCTGCCT TCTGCTTCCT TGTTCAATGG AAACTCGTGA GTACCACCAA GGCCCTCTTC CACATCATCC
1201 CTCCTCTAAT CAAGAACTAT CCATTTCCAT TCACAGTCCC TTCCCCTAAA AACCTCCTGA AATTCTTTTT GCTTTTTAAT TAACAATTAT ATATATTTTA
1301 TCATTACATG ATTCTAGGTG ATCTTTTTTC CAACTGTAAA GCAACCACTG TTACTAAACT GAGGTGTGCT GTCATTTAGA ATAGTATTCC TCTTATTACT
1401 GAACTGTTAG TATCTCAAAT CCTAATCGTC AGTGACATTC TCAGGGGTTG AAACTATTTA AGCACTTTGA AAATAAGCAT TTAAATGGGT GGTATCCTTG
1501 GTTGGAAATG TGTAGCCTAT GGTTTAACCT TGAGATAGAT TTACCTGACT GGCACATCAG AACCCTGTGG ATGTTTGTGT CCTAAAACAA TACTCTTGTT
1601 AGGATCAAAG ACTTTGGATT AAACTACCAG TAGAACTTCT TATTTGATGG CCAAAATAAC CATATAATTA GGATTATATT GCTGAGTGAC TTATTTGATC
1701 AGAAAAATAA ATTTAAGGTG GTAAAAATAA AAAGCAAGCT AGTGATATTT TTTGTATACT GTGATATAGC AATGACTGGC TAGGTTTTGG GGGGCAAAAA
1801 TCTCTAATTA TAGGAAAAGG TAGAGAATGA TTCCTAAAGA TAGCCTTTGG CACTTGCTCA AGTGTAACAT AGGTGAATTA TACAAGTTTG AATTTCACAG
1901 TCACATTATA AATCATCAGT TAAAAAGTAC TCCAAAATAA AAACCGGGTA TTGAGTTATT TTTATGGACT AAAAATGTTA ATGAAAAAAT CATGATAAAT
2001 TTCTATGAAA ATAACAATTA TTCCTCCCGG ATTTCCATTT AAAAAAATAG AAACTTTCTA GAACTTGCAT GCTGGTTGCT TAGTGAAGGG AGTACACGCC
2101 ATTCAAACCT GCCCTGGCAG GTTAATCACT GCTGCAGCCG ATGATTGACA AGCTAATAGT TGTCTTTCCA CTCTGCTGTG TGGCAAACTG TGACACTGCC
2201 ACACTGTCAA CCTGACATCT CAGACTGTTG AAACATGAAT GGTAATGGTG ACTGTGATGA AACTCCAGTT ATAGACAATG ACATCTGTCG TCAGCTGTGC
2301 GGAACCCCTT CCTTTCCAGG AAGTATATGG TGCTGGTCCC GTCCCAGCTC TACACCGACA CCCCTGAGAA AATCTGCCTG CATCTATACC ATCTGAATGA
2401 GACAGTGACT GTGACAGCTT CCTTGATATC TCAAAGGGGA ACGAGAAAAC TGTTCGATGA GCTTGTGGTT GACAAGGACT TGTTCCACTG TCTTTCCTTC  Ex2
2501 ACTGTGAGTA CTAAGTACCC GTCCTGATTT GTTTACATAC TTATGCACAT GCGTGTGAGT ATGTACATAC AATACCCATA CATGTTTCCA AAAAAAGGTT
2601 AATCTAAAAA TATATCTGTG AACTCAAAAC AATAGGAATT TAGAGCATTT ATCCAAAAAA CAAATAGAAT ACATCATCTC ACCCCGAACT GAGAAGGTTC
2701 AATACTTGTC CCATTCAATT CCCTCAAAAA AGCCCACATC CCAGAGAAAG AAAGGGTTCT CCAGTTCATA GCACTTCCTG GTGCTGAGTT AGTAAACTTT
2801 CAAACAGCAC TGATAAAATT AGGACCACAG CTCCTATGCA TAATCCTGAG ATCCTGAATA AGGGCTGGGT TTCTTCTTCA GGACCTCTCT TCAAGACATA
2901 ATTCCTATCA CCATTAAAAC ATTACCATTC TCTCTACCTC AGTTTGAGAG TTAGAGCTAA GTTTCTCTCA TGTCTTAGAG TTAATTTTCA ATGAAGATTA
3001 TATTTACTCT GTAAAGATGA CTGTGTTTTT GTTGTTCACA TAAACTCCCA CTGACATACA GAAAATGCTT TTTCTTTTCA TCCCTTTCTT TTCCTGACAA
3101 TTTTTCCCTC CATGTAATAC TTAAAGAGAA ACTGATGGAG CCTCCACCTC TTTCCTGGGA GGTGGAAGGA GAAAAGGTTA ATGAGGTAAT GTAAGAAAGG
3201 CAGAGGGAGA TGGTGAAGGG TCTGTGTTCC TGTCGTTCTC AGATCCCTAG ATTGCCCTCT TCTGAAGAGG AGGAGTCACT CGACATCAAC ATAGAAGGGG
3301 CAAAGCATAA ATTCAGCAAA AGGCGTGTGG TGCTTGTGAA GAACAAAGAA AGTGTTGTCT TTGTCCAGAC TGATAAGCCC GTGTACAAGC CAGGACAGTC  Ex3
3401 AGGTATGAAG CTCCACGGTG GGCAAAGGAA TTAATACACC AACCTAAACA TGGTGTGAAG AAGGCCTGGG TCTGTGGACC TCACTAGAAA ATTCTTTCAC
3501 TTTTTCCAGT TAAATTTCGG ATTGTCTCTA TGGATAAAAA TCTACACCCC CTGAATGAGT TGGTGAGTTT TTTACACGTG GGTATTGGTT GGTTCTTCAA
3601 GACAAAATGT CAGAGCAATA TTCTAATTTC CAGTTCCTTC TGGGACATGA GTTAGCCTTT TTTTTCTATT CTGTTCTTTT GCCTATTTAT ACCCCTTACT
3701 CTTTTACCTT TGGCTTCCTT ATATGTGACT GTTCTTCTTT GGACTCCATT ACTGGATAAG TATCTCTGGG TTTTCTTCAT ACTCTGTGGA GACTTATACT
3801 GTATAGAACC TTTTAGGTTT TTCAGTCATT TCTTTTGCAG ATTCTGTTAC TTTGTGCCAT CCCTCAGAAA TCCATTCCTA GTGTAATAGG GATATTTCCC
3901 ACACAGTCTC CATTATTTTT ATAACCTCGT ATTTAACACT TTGGGTCTTC CCTAGATGGA CTCTTCTAAG AATAAAATTG CATTTCAGTC TCTTGTGCTC
4001 TGTGCTTAGG CAAGTTGTGT GTATCTGTGT TAAATACTCT CTCCTATGAA TAGAAGGGTC TCTGATGAGG GATGAGGGAT GCTTTAATCT AG
```

FIGURE 3: DNA sequence of the 5′-terminal regions of the α_1I3 and α_1I4 genes. (A) Sequence of the 5′-terminal region of the α_1I3 gene. Exons 1, 2, and 3 are underlined and designated by Ex1, Ex2, and Ex3 on the right. The $TATA_M$ box (see Figure 6) is underlined, and the transcription start site is indicated by an arrow. The signal peptide oligonucleotide sequence used as a hybridization probe (Figure 1B) and as a primer (Figure 6) is underlined in exon 1. (B) Sequence of the 5′-terminal region of the α_1I4 gene. Exons 1, 2, and 3 are labeled as in panel A and underlined. The sequence qualifying as a TATA box is underlined; however, the transcription start site is not defined, because this gene is not expressed in liver (see the text). The end of exon 1 is defined only by analogy with exon 1 of the α_1I3 gene.

approximately 2–3-fold reduced concentrations in acute-phase livers, 18 h after experimental induction of an inflammatory response (Figure 5, track 6). No mRNA species derived from the α_1I4 gene were detected in normal liver (Figure 5, tracks 2 and 5). The M13 subclones of the α_1I4 gene have been verified by DNA sequence analysis and were confirmed to carry the appropriate inserts. Therefore, we conclude that the α_1I4 gene is not transcribed in liver to an extent detectable within the sensitivity range of this technique.

Primer extension experiments were performed using a 30-nucleotide synthetic oligonucleotide primer from the signal peptide coding sequence of the prototype α_1I3 gene (Figure 6). Extension was carried out either in the absence (Figure 6, track 1) or in the presence of dideoxynucleotide triphosphates (Figure 6, tracks 2–5). The first reaction produced full-length cDNA transcripts of the region between the primer

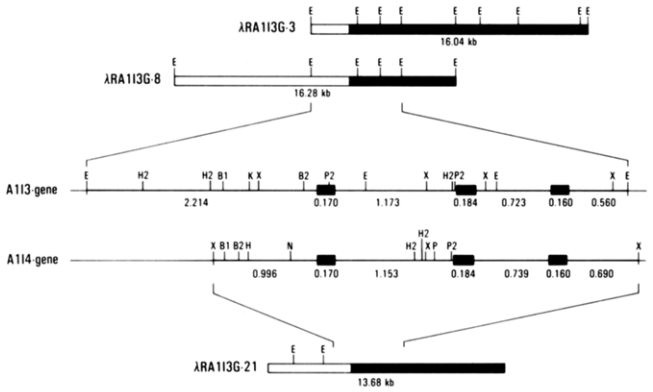

FIGURE 4: Schematic comparison of the 5′-terminal regions of the α_1I3 and α_1I4 genes; restriction maps. The inserts of the two group I clones λRA1I3G-3 and -8, representing the 5′-terminal region α_1I3 gene, and that of group III clone λRA1I3G-21, representing the 5′-terminal region of the α_1I4 gene, are given on top and at the bottom. The boxes represent the entire insert carried in these clones; the length in kilobase pairs is indicated below the box. E: cleavage sites for the restriction enzyme *Eco*RI. The open part of the boxes represents 5′-flanking sequences and the closed part the transcribed portion of the α_1I3 gene and the analogous portion of the α_1I4 gene. The blownup portion in the middle is a restriction map for the sequenced regions (Figure 3A,B). The black boxes represent the exons; the connecting line represents the introns and the 5′-flanking sequence. The sizes of the exons and introns are indicated below (in kbp). The 5′- boundary of exon 1 of the α_1I4 gene is defined only by analogy with the α_1I3 gene. B1, *Bgl*I; B2, *Bgl*II; X, *Xba*I; K, *Kpn*I; P2, *Pvu*II; P, *Pst*I; H, *Hin*dIII; H2, *Hpa*II; N, *Nco*I.

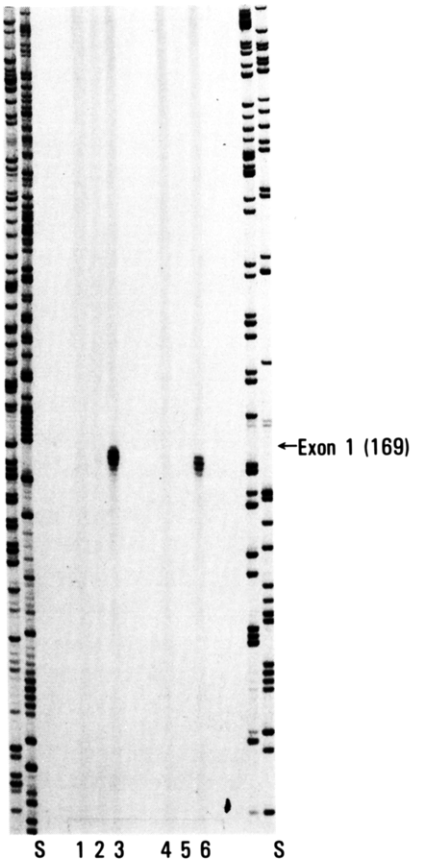

FIGURE 5: Mapping of the transcription start site of the α_1I3 gene by S1 analysis. Internally labeled single-stranded phage DNA corresponding to the coding strands in the 5′-end region of the α_1I4 gene (tracks 2 and 5) and the α_1I3 gene, respectively (tracks 3 and 6), was annealed with rat liver mRNA from normal rats (tracks 1–3) and 18 h after experimental induction of an inflammatory response (tracks 4–6). In tracks 1 and 3, single-stranded phage DNAs representing the noncoding strands of the α_1I4 and α_1I3 genes were used respectively as negative controls. The hybrids were treated with S1 nuclease, and S1-resistant fragments were electrophoretically separated in a 6% polyacrylamide–6 M urea gel. The two tracks labeled S on the right and left are a sequencing ladder of an unrelated, known DNA fragment, serving as a size marker. The arrow points to the top band of the multiplet in track 3, which defines the size of exon 1 of the α_1I3 gene (169 nucleotides).

and the 5′ end of the message. From the length of these products, the position of the cap site was deduced. The second reaction produced sequence information of the RNA region located between the primer and the cap site. This sequence was identical with the genomic DNA sequence of clone λRA1I3G-3, providing additional support for the correct location of the cap site in the gene. The cap site as mapped by this technique (designated M in Figure 6) was located within ±4 nucleotides of the position determined by S1 mapping. Therefore, we conclude that position M on the genomic sequence (Figure 6) corresponds to the cap site of the α_1I3 mRNA and thus to the transcription start site of the gene. Among the primer extension products, an abundant group of transcripts 39 nucleotides shorter than the major product was detected (Figure 6). These could reflect either strong stops, due to pausing of the polymerase (caused, for example, by a strong secondary structure of the RNA), or transcripts originating from a second, minor start site, designated m in Figure 6. The S1 mapping experiment (Figure 5) did not reveal transcripts originating from a minor start site, and therefore, these fragments are probably not due to the utilization of a minor start site. A TATA box motif AAATAA, designated $TATA_M$ in Figure 6, is present 17 nucleotides 5′ of the cap site in the α_1I3 gene.

Transcription of the α_1I3 Gene Is Decreased during Acute Inflammation. Hepatic α_1I3 mRNA concentrations are reduced 3–10-fold below normal levels during an acute-phase response (Aiello et al., 1988; Braciak et al., 1988; Lamri et al., 1988). To determine whether this effect is due to reduced transcription rates, nuclear run-on experiments were performed (Figure 7A). An acute-phase response was experimentally induced by intraperitoneal injection of complete Freund's adjuvant (CFA) into a series of adult rats. At various times after injection, livers were excised, nuclei were prepared, and transcription rates of the α_1I3, α_2M, and rat serum albumin (RSA) genes were measured (Figure 7A and Experimental Procedures). The transcription rate of the α_1I3 gene was surprisingly high in normal rats and was decreased with a minimum reached at 6–9 h after inflammation. It recovered slowly from 12 h on. At 24 h, it had returned to about one-third of the normal rate. This experiment was performed 3 times, and the factor of reduction at 6 h ranged from 10-fold to 19-fold with an average of 12.7-fold. At the same time, the transcription rates of the RSA gene were transiently reduced between 2- and 3-fold, as previously described (Figure 7A; Gehring et al., 1987; Schreiber, 1987). The α_2M rates were increased on average 5.3-fold, with a maximum of 8-fold at 12 h. We conclude that the reduced α_1I3 mRNA levels observed during an acute-phase response are due to a pronounced transient decrease of the transcription rate of the α_1I3 gene.

Nuclear α_1I3 RNA precursor levels were measured during the time course of the acute inflammatory response (Figure 7B). Nuclei from the same preparations were used as those employed in the nuclear run-on experiments of Figure 7A. Nuclear RNA was prepared, deposited on a nylon membrane, and quantitated by hybridization with an α_1I3 gene intron 1 specific probe. This experiment revealed a very rapid decrease in nuclear α_1I3 precursor RNAs, as early as 2 h after injection (Figure 7B). The precursor RNA levels were lowest at 6 h

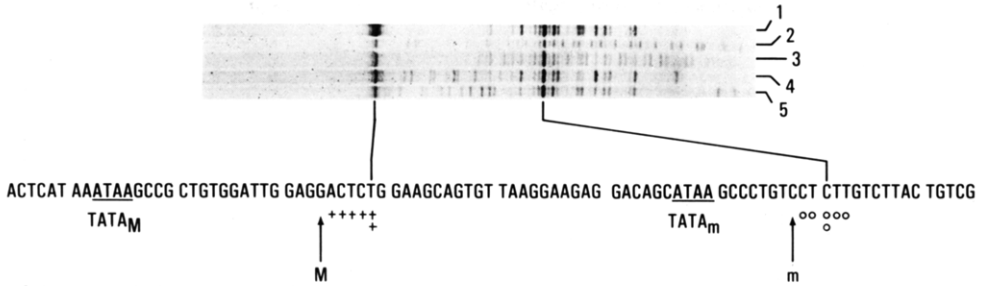

FIGURE 6: Mapping of the transcription start site of the α_1I3 gene by primer extension/RNA sequencing. Size-enriched, polyadenylated mRNA was annealed with the signal peptide oligonucleotide primer (Figure 3A). Primer extension was carried out with reverse transcriptase as described under Experimental Procedures, either in the absence (track 1) or in the presence of dideoxynucleotide triphosphates (tracks 2–5). The main band in track 1 represents transcripts extending to the T residue indicated in the genomic sequence (bottom). Four additional weaker bands of larger size were present on the autoradiograph, which were indicated by (+) symbols in the sequence. The transcription start site was therefore placed at the G residue, labeled by the arrow and marked M. A group of intense transcripts 39 nucleotides shorter than the full-length transcripts is visible in all five tracks. The strongest of these bands corresponds to the C residue indicated by two open circles in the sequence. A potential minor transcriptional start site may be located at the C residue marked by the arrow and labeled m. Tracks 2–5 contained respectively ddATP, ddGTP, ddCTP, and ddTTP.

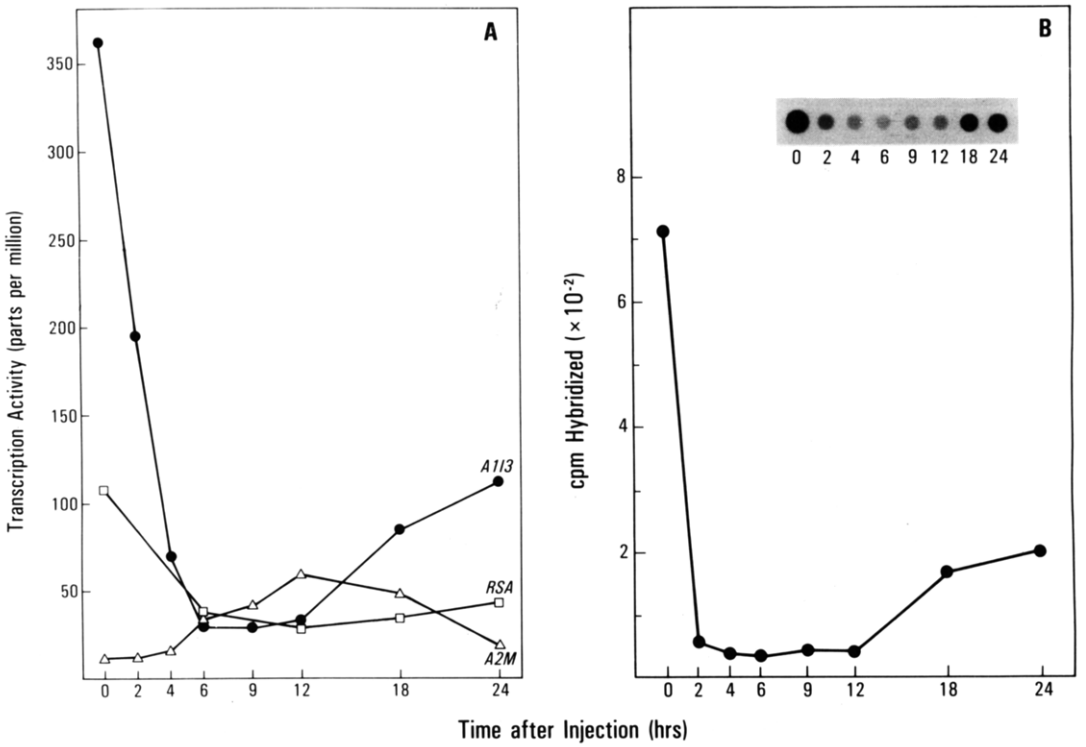

FIGURE 7: Transcription rates and nuclear RNA levels of the A1I3 gene during acute inflammation. Acute inflammation was provoked in adult male rats by intraperitoneal injection of complete Freund's adjuvant (CFA) as described (Gehring et al., 1987). Nuclei were prepared from livers excised at various times after injection, and transcription rates were measured by published procedures (Northemann et al., 1985). (A) (50–100) × 10^6 cpm (Cerenkov counts) of in vitro elongated RNA was used as input in each hybridization reaction. The transcription rates are expressed, in parts per million, as the fraction of the input material that hybridized to cDNA probes immobilized on filters. A1I3 (●), α_1I3; RSA (□), rat serum albumin; (△), α_2M; (B) Nuclear RNA prepared from purified nuclei isolated at various times after inflammation was deposited on nylon membrane as detailed under Experimental Procedures. The filter was hybridized with a nick-translated DNA probe of 850 bp, representing intron 1 of the α_1I3 gene. Precursor RNA levels are expressed as counts per minute hybridized. Quantitation was performed by excision of the dots after autoradiography and quantitative liquid scintillation counting in a commercial scintillation cocktail, using a Beckman Model LS8000 β counter.

with an average 16-fold reduction. This experiment supports the conclusion of rapidly reduced transcription rates of the α_1I3 gene during an acute-phase response.

Glucocorticoids Alone Are Sufficient To Cause Decreased α_1I3 Transcription in Rats. A number of other rat genes are controlled during the acute-phase response by two categories of signal substances, glucocorticoids and monokines of the hepatocyte stimulating factor type [HSF; for reviews, see Gordon and Koj (1985), Schreiber (1987), and Fey and Fuller (1987)]. Glucocorticoids alone or unable to produce a major transcription change of some of these genes, including the α_2M gene, and achieve a maximal effect on these genes only in combination with other inflammatory mediators (Northemann et al., 1988a). In contrast, glucocorticoids alone are sufficient to cause strong transcriptional changes of a second group of genes, including the α_1-acid glycoprotein gene (α_1AGP; Kulkarni et al., 1985; Baumann & Maquat, 1986; Klein et al., 1987). Therefore, we wanted to determine whether glucocorticoids alone are able to affect the transcription of the α_1I3 gene in rats and to compare the magnitude of this effect with the changes occuring during an acute inflammatory response. A high dose of the synthetic glucocorticoid dexamethasone was injected into rats, and the transcription rates of the α_1I3 gene were measured by nuclear run-on experiments at several times after injection (Figure 8A). The experiment was performed twice with very similar results, and the tran-

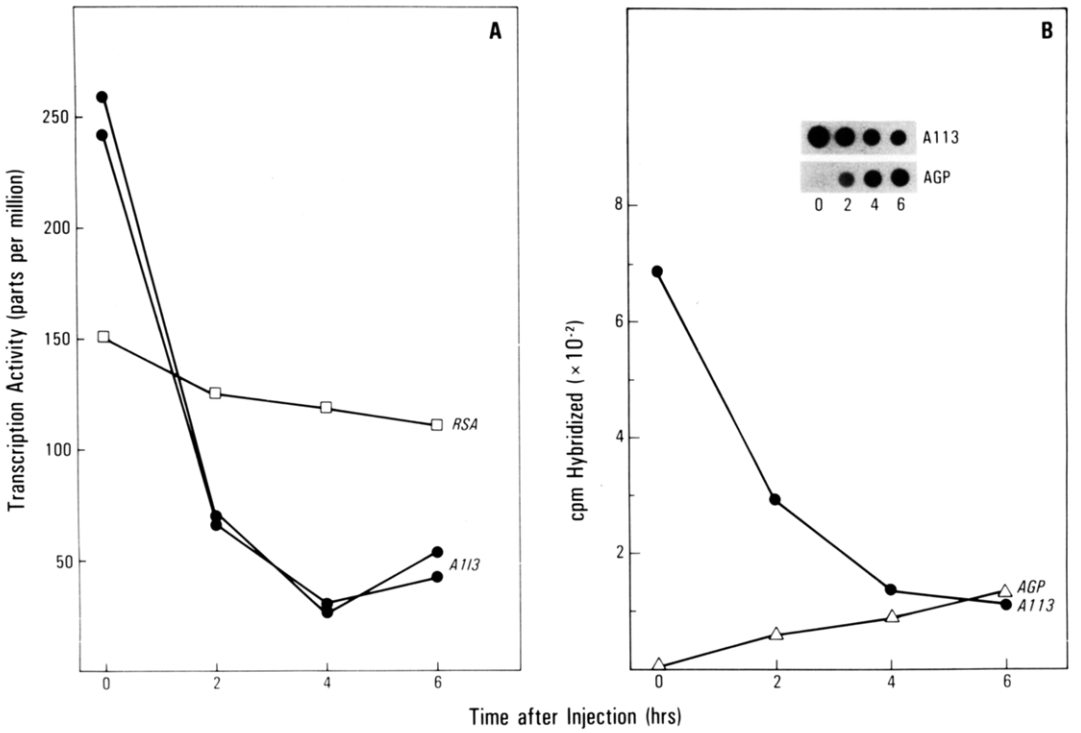

FIGURE 8: Transcription rates and nuclear RNA levels of the α_1I3 gene after injection of glucocorticoids into rats. Dexamethasone (Hexadrol, a 4 mg/mL solution of dexamethasone sodium phosphate) was injected intraperitoneally into rats at a concentration of 4 mg/kg body weight, i.e., 0.25 mL into 250-g adult male rats. At various times after injection, livers were excised, and nuclei were prepared. Transcription rates were measured by nuclear run-on experiments, and nuclear precursor RNA levels were measured by dot blot hybridization as described for Figure 7 and under Experimental Procedures. (A) A1I3 (●), α_1I3; RSA (□), rat serum albumin. The two curves represent two independent experiments. (B) A1I3 (●), α_1I3; AGP (△), α_1-acid glycoprotein. The insert shows the autoradiographic evaluation of the dot blots; the curves show the quantitation of the same dots by liquid scintillation counting.

scription rate was reduced on average 9.1-fold 4 h after injection. RSA gene transcription was reduced by about 20%.

Nuclear α_1I3 precursor RNA levels were measured with nuclei from the same time points as in Figure 8A and were quantitated by hybridization with α_1I3 and α_1AGP gene-specific intron probes (Figure 8B). The α_1I3 nuclear precursor RNA concentration was decreased 6.2-fold at 6 h, while at the same time the α_1AGP precursor concentration was increased 39.5-fold.

From these two experiments together, we conclude that high doses of glucocorticoids (dexamethasone) alone are sufficient to reduce transcription of the α_1I3 gene in rats to approximately half the extent which occurs in a fully developed acute-phase response.

Discussion

The identity of the isolated genomic DNA clones was confirmed by the perfect match between the sequence of the first three exons of group I clones with the published prototype α_1I3 DNA sequence (Braciak et al., 1988). The existence of an α_1I3 gene family was established by the genomic Southern blot result and the isolation of five groups of related but structurally distinct α_1I3 genomic clones. It is difficult to evaluate the precise extent of this gene family, because we may not yet have isolated clones representing all members of this family, and because some of the bands on a Southern blot may represent multiple genes. From the available data, we conclude that this family likely contains four or more members. It is difficult to discriminate between different gene loci and alleles at one locus. Although inbred rats were used to prepare the DNA for the Southern blots and the genomic DNA libraries, we cannot be certain that these rats were homozygous at all loci. However, from the results presented here, it is possible to conclude that the rat genome counts at least two, and possibly more, α_1I3 gene loci. In contrast, α_2M and the related complement C3 gene are single-copy genes in humans and rats (Whitehead et al., 1982; Davies et al., 1983; Kan et al., 1985) while complement C4, a member of the same gene family, forms a small subfamily with two or three loci in the murine and human genomes (Carroll et al., 1984). Within the sensitivity limits of the techniques employed in this study, we have not detected any transcripts of the α_1I4 gene in rat livers, compatible with the interpretation that the α_1I4 gene could be a nontranscribed pseudogene. However, we have not excluded the possibility that it could be transcribed in nonhepatic cell types. The α_1I4 gene will provide a useful negative control for further studies of the control elements of liver-specific expression of the α_1I3 gene. The 5′-flanking region of the α_1I4 gene is over 90% conserved in the α_1I3 gene for the first 1 kbp, and yet this gene is not expressed in liver. Consequently, the sequence differences between the α_1I3 and α_1I4 genes should provide helpful indications to locate sequences essential for the hepatic expression of the α_1I3 gene.

The α_2M protein has an approximate molecular weight of 180 000, and the α_2M gene is 48 kbp in length (Northemann et al., 1988). The α_1I3 protein is of comparable size as α_2M, and both genes have evolved from a common ancestor. Therefore, the α_1I3 gene is likely to be of similar size and organization as the α_2M gene. Our interest is focused on the regulatory regions of the α_1I3 gene. Although regulatory elements may be located both in the 5′ and in the 3′ portions of the gene and even internally, we have concentrated here on the 5′ part first, because it will permit us to compare the 5′-terminal control elements of this gene with those of many others.

As the relative half-life of intron-containing nuclear RNA transcripts is short (Nevins & Darnell, 1986; Granner et al., 1983), we reasoned that measurement of their levels should

FIGURE 9: Conserved glucocorticoid receptor binding elements and glucocorticoid mediator elements in the 5′ control region of the α_1I3 gene. Sequence alignments were performed with the Bestfit algorithm (Devereux et al., 1984). (a) The α_{2u}-globulin sequence was from Addison and Kurtz (1986). (b) The α_1FP sequence was from Belanger et al. (1987). (c) The α_1I3 gene sequence was from Figure 3A. (d) The α_1AGP sequence was from Baumann and Maquat (1987) and Reinke and Feigelson (1985). The consensus GRE sequence was from Jantzen et al. (1987). Negative numbers preceding the sequence lines indicate the nucleotide positions of the first listed character in the original sequence. For the α_1I3 gene, nucleotide −210 corresponds to nucleotide 2004 in Figure 3A. The two boxed sequences in the left part of the figure include the conserved GRE site of the α_1I3 and α_1FP genes. The hexanucleotide core GRE is underlined. The box in the right part represents the inverted GRE element, abbreviated GME. The underlined consensus hexanucleotide sequence GG$^{A}_{C}$A$^{C}_{A}$A includes the GGAACA sequence found by Addison and Kurtz (1986) and Klein et al. (1987) to be functionally relevant in the α_{2u}-globulin and α_1AGP genes. Dashes in the α_1I3 sequence: a gap introduced to optimize the alignment with the α_1FP gene.

reflect changes in transcription rates more closely than the titration of cytoplasmic mRNA levels. Indeed, the kinetics of the intranuclear α_1I3 precursor RNA levels during the inflammatory response followed the kinetics of the transcription rates, as independently measured, very closely (Figure 7B). This result supports the transcription rate measurements reported above. The maximal decrease in nuclear precursor RNA concentration (16-fold) was comparable to the reduction of the transcription rate (12.7-fold). Unexpectedly, the precursor RNA level was reduced more rapidly than the transcription rate (Figure 7), suggesting an element of posttranscriptional control may be operative in the overall regulation of α_1I3 gene expression. However, we conclude that the regulation of the α_1I3 gene during an acute-phase response occurs primarily through a rapid and transient reduction of its transcription.

Many genes are jointly regulated during an acute-phase response by glucocorticoids and by other inflammatory mediators including the monokines interleukin 1 (IL-1), tumor necrosis factor (TNF), and hepatocyte stimulating factor/interleukin 6 [HSF/IL6; for reviews, see Gordon and Koj (1985), Schreiber (1987), Fey and Fuller (1987), and Gauldie et al. (1987)]. Some acute-phase genes, such as the α_1-acid glycoprotein gene in rats (α_1AGP), can be regulated, at least partially, by glucocorticoids alone at the transcriptional level (Kulkarni et al., 1985a,b; Baumann & Maquat, 1986; Klein et al., 1987). Therefore, we have asked whether glucocorticoids alone can affect the transcription of the α_1I3 gene or whether the combined action with other mediators is required to produce this effect. To dissect the effects of these two categories of signal substances, we have compared the effects of an injection of glucocorticoids alone into rats on the transcription of the α_1I3 gene with the changes occurring in a full acute-phase response, where both agents are known to be active. Figure 8A,B demonstrates that glucocorticoids can act independently on the α_1I3 gene: injection of dexamethasone alone into healthy, adult male rats produced a decrease of the α_1I3 transcription rate of approximately half the extent observed in a full acute-phase response. Thus, the α_1I3 gene responds to glucocorticoids in a similar mode as the α_1AGP gene. From these results, we extrapolate that in an acute-phase response the effects of both types of signals may be additive and that both glucocorticoids and other inflammatory mediators may act independently rather than synergistically on the α_1I3 gene.

The nuclear α_1I3 precursor RNA levels, titrated with an intron probe, declined after injection of dexamethasone into rats with similar kinetics as the transcription rates (Figure 8A,B), supporting the transcription rate measurements.

Apart from the acute-phase genes, many other genes expressed in rodent hepatocytes are regulated by glucocorticoids, and a variety of different response patterns have been observed. Some genes are transcriptionally induced by glucocorticoids (α_1AGP and α_{2u}-globulin; Kulkarni et al., 1985a,b; Baumann & Maquat, 1986; Klein et al., 1987; Addison & Kurtz, 1986); others are decreased in their transcription [α_1-fetoprotein (α_1FP) and procollagen; Turcotte et al., 1985; Weiner et al., 1987]. The glucocorticoid effect on some of these genes is blocked by protein synthesis inhibitors (α_1AGP and α_{2u}-globulin; Klein et al., 1987, 1988; Addison & Kurtz, 1986); for others, it is not (α_1FP; Turcotte et al., 1985). We have compared the 5′-flanking sequences of the α_1I3 gene by computer-assisted sequence analysis with those of the α_{2u}-globulin, α_1FP, and α_1AGP genes as representatives of these different categories of regulatory responses to glucocorticoids.

The hexanucleotide 5′TGTYCT3′ is commonly referred to as the GRE (glucocorticoid responsive element) core sequence. It is the common core of two different extended versions of the consensus glucocorticoid receptor binding sequence (Jantzen et al., 1987; Scheidereit et al., 1983; Payvar et al., 1983; Karin et al., 1984; Moore et al., 1985). The glucocorticoid hormone–receptor complex can bind to different DNA target sites of approximately 16–20 bp in length that contain this core.

The 5184 bp sequence of the 5′-terminal region of the α_1I3 gene (Figure 3A) contains 20 copies of the hexanucleotide GRE core element, 12 in forward and 8 in reverse orientation. Most of these do not share significant similarity with the extended GRE consensus sequences and therefore are not likely to be functionally relevant. However, two of these elements, located at −158 to −163 bp and −197 to −202 bp, are likely candidates for functionally relevant GRE elements for the following reasons (Figure 9). The more distal element at −197 to −202 bp shares a 14/16-nucleotide sequence identity with the region around position −166 of the α_1FP gene, which is a confirmed glucocorticoid receptor binding site (Belanger et al., personal communication; Chevrette et al., 1987). The other element at −158 to −163 bp is an imperfect copy of the core sequence, congruent in only 5/6 positions. It is present in reverse orientation in the α_1I3 gene. However, the area of the α_1I3 gene surrounding this element shares extended sequence conservation with the region −114 to −121 of the α_1AGP gene, and with the region −118 to −125 of the α_{2u}-globulin gene. These areas have been shown by transfection experiments to be relevant for glucocorticoid control of the α_1AGP and α_{2u}-globulin genes (Baumann & Maquat, 1986; Addison & Kurtz, 1986). The extent of sequence similarity with the α_1AGP gene around this site is particularly striking, raising the possibility that both genes share in this position binding sites not only for the glucocorticoid receptor but possibly for additional other proteins binding in the vicinity that could modulate the glucocorticoid control of these genes (Klein et al., 1988).

It will be interesting to determine in the future how glucocorticoids down-regulate the transcription of this gene and of a few other genes, while in most known cases glucocorticoids increase the transcription of responsive genes.

ACKNOWLEDGMENTS

We thank Dr. T. Sargent for the rat genomic DNA libraries and Drs. K. Lonberg-Holm and L. Abraham for critical reading of the manuscript. We are grateful to Drs. Belanger, Feigelson, and Kurtz for communication of unpublished data. The help of Geoffrey Hudson with the synthesis of oligonucleotides, Gene Jensen with the VAX computer, and Keith Dunn with the preparation of the manuscript is sincerely acknowledged.

Registry No. α1I3, 88943-21-9; dexamethasone, 50-02-2.

REFERENCES

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

Aiello, L. P., Shia, M. A., Pilch, P. F., & Farmer, S. R. (1988) *J. Biol. Chem. 263*, 4013–4022.

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.

Barrett, A. J. (1981) *Methods Enzymol. 80*, 737–753.

Baumann, H., & Maquat, L. E. (1986) *Mol. Cell. Biol. 6*, 2551–2561.

Belanger, L., Guertin, M., LaRue, H., Turcotte, B., Chevrette, M., & Wrange, O. (1987) *Abstracts, Cold Spring Harbor Meeting on the Regulation of Liver Gene Expression*, Cold Spring Harbor Laboratory, New York.

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

Carroll, M. C., Campbell, R. D., Bentley, D. R., & Porter, R. R. (1984) *Nature (London) 307*, 237–241.

Chevrette, M., Guertin, M., Turcotte, B., & Belanger, L. (1987) *Nucleic Acids Res. 15*, 1338–1339.

Davies, K. E., Jackson, J., Williamson, R., Harper, P. S., Ball, S., Sarfarazi, M., Meredith, L., & Fey, G. (1983) *J. Med. Genet. 20*, 259–263.

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

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

Esnard, F., Gutman, N., El Moujahed, A. E., & Gauthier, F. (1985) *FEBS Lett. 182*, 125–129.

Fey, G. H., & Fuller, G. (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., & Koj, A., Eds. (1985) in *The Acute Phase Response to Injury and Infection*, Elsevier, Amsterdam.

Granner, D., Andreone, T., Sasaki, K., & Beale, E. (1983) *Nature (London) 305*, 549–551.

Huang, J. S., Huang, S. S., & Denel, T. F. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 342–346.

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.

Klein, E. S., Reinke, R., Feigelson, P., & Ringold, G. M. (1987) *J. Biol. Chem. 262*, 520–523.

Klein, E. S., DiLorenzo, D., Posseckert, G., Beato, M., & Ringold, G. (1988) *Mol. Endocrinol.* (in press).

Kulkarni, A. B., Reinke, R., & Feigelson, P. (1985a) *J. Biol. Chem. 260*, 15386–15389.

Kulkarni, A. B., Gubits, R. M., & Feigelson, P. (1985b) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2579–2582.

Lamri, Y., Giuily, N., Legres, L., Bernuau, D., Fey, G., & Feldmann, G. (1988) *J. Cell Biol.* (submitted for publication).

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.

Nevins, J. R., & Darnell, J. E., Jr. (1978) *Cell (Cambridge, Mass.) 15*, 1477–1493.

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

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

Northemann, W., Shiels, B. R., Braciak, T. A., Hanson, R. W., Heinrich, P. C., & Fey, G. H. (1988b) *Biochemistry 27*, 9194–9203.

Payvar, F., Defranco, D., Firestone, G. L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J. A., & Yamamoto, K. R. (1983) *Cell (Cambridge, Mass.) 35*, 381–392.

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.

Scheidereit, C., Geisse, S., Westphal, H. M., & Beato, M. (1983) *Nature (London) 304*, 749–752.

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

Schweizer, M., Takabayashi, K., Geiger, T., Laux, T., Biermann, G., Buhler, J. M., Gauthier, F., Roberts, L. M., & Heinrich, P. C. (1987) *Eur. J. Biochem. 164*, 375–381.

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

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

Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Lonblad, P. B., Jones, C. M., Wierzbicki, D. M., Magnusson, S., Domdey, H., Wetsel, R. A., Lundwall, A., 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.

Turcotte, B., Guertin, M., Chevrette, M., & Bélanger, L. (1985) *Nucleic Acids Res. 13*, 2387–2398.
Weiner, F. R., Czaja, M. J., Jefferson, D. M., Giambrone, M.-A., Kaspa, R. T., Reid, L. M., & Zern, M. A. (1987) *J. Biol. Chem. 262*, 6955–6958.
Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W. F., Povey, S., & Fey, G. H. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5021–5025.

Heptads of Polar Ether Lipids of an Archaebacterium, *Methanobacterium thermoautotrophicum*: Structure and Biosynthetic Relationship†

Masateru Nishihara, Hiroyuki Morii, and Yosuke Koga*
Department of Chemistry, University of Occupational and Environmental Health, Japan, Yahatanishi-ku, Kitakyushu 807, Japan

Received May 16, 1988; Revised Manuscript Received July 29, 1988

ABSTRACT: The structures of the eight major polar lipids of *Methanobacterium thermoautotrophicum* were determined. They were one diether glycolipid (gentiobiosylarchaeol) and serine-, inositol-, and ethanolamine-containing diether and tetraether types of phospholipids and phosphoglycolipids [archaetidyl-L-serine, caldarchaetidyl-L-serine, gentiobiosylcaldarchaetidyl-L-serine, 1D-1-(archaetidyl)-*myo*-inositol, 1D-1-(caldarchaetidyl)-*myo*-inositol, 1D-1-(gentiobiosylcaldarchaetidyl)-*myo*-inositol, archaetidylethanolamine]. In combination with 2 neutral lipids and 3 polar lipids that have been already described in the previous paper [Nishihara, M., Morii, H., & Koga, Y. (1987) *J. Biochem. (Tokyo) 101*, 1007], the 13 lipids were proposed to be classified in three groups, that is, three "heptads", each of which was constituted by diether and tetraether types of neutral lipids, glycolipids, and phospholipids and a tetraether phosphoglycolipid. The heptad concept implied the biosynthetic relationship between diether and tetraether lipids which was supported by in vivo kinetic experiments. When growing cells were pulse labeled with [^{32}P]orthophosphate, there was a lag of 15–90 min between the rapid incorporation of label into diether polar lipids and that of label into the corresponding tetraether polar lipids. The lag times and rates of incorporation of ^{32}P into tetraether phospholipids and their respective diglucosyl derivatives (phosphoglycolipids) were almost identical. In a pulse–chase experiment with [^{32}P]P_i, rapid turnover of the three diether lipids other than archaetidylethanolamine was observed. At the same time radioactivity was incorporated into gentiobiosylcaldarchaetidylinositol and other tetraether polar lipids. These results are consistent with a model which postulates that head-to-head condensation of phytanyl chains of two diether polar lipids occurs to yield tetraether polar lipids.

One of the prominent distinguishing features of archaebacteria is the presence of glycerol isopranyl ether lipids. The complex lipids are classified as diether and tetraether types of lipids by their hydrophobic core portion. Because the tetraether types of lipids are apparently made of two halves of diether types of lipids in structure, it is important to elucidate the structural and biosynthetic relationship of the two types of lipids in one archaebacterium. For this purpose, structure determination of a whole set of major polar lipids in a methanogen is desirable, because, in contrast with extreme halophiles and sulfur-dependent archaebacteria, methanogens have both types of lipids (Balch et al., 1979; Tornabene & Langworthy 1979). Although several polar lipids have been reported in various methanogenic bacteria, only one report on the structure analysis of the major polar lipids in one species (*Methanospirillum hungatei*; Kushuwaha et al., 1981) has appeared.

Recently, the lipid composition and structures of three tetraether polar lipids of *Methanobacterium thermoautotrophicum* were reported (Nishihara & Koga, 1987; Nishihara et al., 1987). These lipids were caldarchaetidylethanolamine, gentiobiosylcaldarchaeol, and gentiobiosylcaldarchaetidylethanolamine. We have proposed in that paper the concept of "a quartet of lipids" which consisted of the three lipids along with bare caldarchaeol. Moreover, it has been suggested that a quartet would be extended to a heptad by the addition of three diether lipids. Kramer et al. (1987) have found archaetidylethanolamine, which should be a diether component of the heptad, in *M. thermoautotrophicum*. As previously reported (Nishihara & Koga, 1987), this organism has at least 23 species of polar lipids, only 4 of which were structurally elucidated as described above. In this paper the structures of eight more polar diether and tetraether lipids of *M. thermoautotrophicum* are determined, and three sets of complete heptads of serine, inositol, and ethanolamine lipids are described.

Although the heptad concept is conceived on a structural basis, it implies the biosynthetic relationship of diether and tetraether types of polar lipids. Therefore, the hypothetical relationships should be experimentally examined. Up to now, mechanisms of biosynthesis of tetraether lipid from diether lipids in archaebacteria remain speculative. On the basis of the examination of the structural regularities in the tetraether lipids, Langworthy (1985) and De Rosa and Gambacorta (1986) inferred that the condensation of diether neutral lipid occurred before attachment of polar head groups to the diether

† This study was supported in part by Grants-in-Aid for Scientific Research (62560092) from the Ministry of Education, Science and Culture of Japan and grants from the Agricultural Chemistry Foundation, Japan.

* Address correspondence to this author at the Department of Chemistry, School of Medicine, University of Occupational and Environmental Health, Japan, 1-1 Iseigaoka, Yahatanishi-ku, Kitakyushu, Fukuoka 807, Japan.

0006-2960/89/0428-0095$01.50/0