MOLECULAR AND CELLULAR BIOLOGY, Jan. 1986, p. 302–314
0270-7306/86/010302-13$02.00/0

Vol. 6, No. 1

Rat Metallothionein-1 Structural Gene and Three Pseudogenes, One of Which Contains 5′-Regulatory Sequences

ROBERT D. ANDERSEN,[1]* BRUCE W. BIRREN,[1,3] SUSAN J. TAPLITZ,[1,2] AND HARVEY R. HERSCHMAN[1,2,3]

Laboratory of Biomedical and Environmental Sciences,[1] Department of Biological Chemistry,[2] and Molecular Biology Institute,[3] University of California at Los Angeles, Los Angeles, California, 90024

Received 1 July 1985/Accepted 18 October 1985

As shown by Southern blot analysis, the metallothionein-1 (MT-1) genes in rats comprise a multigene family. We present the sequence of the MT-1 structural gene and compare its features with other metallothionein genes. Three MT-1 pseudogenes which we sequenced apparently arose by reverse transcription of processed mRNA transcripts. Two of these, MT-1ψa and MT-1ψc, are retrogenes which derive from the MT-1 mRNA, having diverged from the MT-1 gene 6.9 and 2.6 million years ago, respectively. The third, MT-1ψb, differs from the MT-1 cDNA by only three nucleotide alterations. Surprisingly, MT-1ψb also preserves sequence homology for 142 base pairs 5′ to the transcription initiation site of the parent gene; it contains a promoter sequence sufficient for specifying metal ion induction. We identified, by S1 nuclease mapping, an RNA polymerase II initiation site 432 base pairs 5′ of the MT-1 transcription initiation site of the MT-1 structural gene which could explain the formation of the mRNA precursor to this pseudogene. We were unable to detect MT-1ψb transcripts, either in liver tissue or after transfection. We conclude that the absence of detectable transcripts from this pseudogene is due to either a reduced level of transcription or the formation of unstable transcripts as a consequence of the lack of a consensus sequence normally found 3′ of transcription termination in the MT-1 structural gene.

Metallothioneins are 6,000-dalton, cysteine-rich proteins that bind the metal ions zinc, copper, cadmium, and mercury (see 32 for review). Metallothioneins are usually isolated as two major isoforms, designated MT-1 and MT-2. The metallothionein (MT) genes are inducible by metal ions (4, 17), glucocorticoid hormones (19, 20, 36, 37), stress (51), bacterial endotoxin (16), iodoacetate (18), and interferon (22). Human MT-1 and MT-2 genes show differential expression with metal ion and hormone inducers (57). The proposed functional roles of metallothioneins include: protection against heavy metal ion intoxication by cadmium or mercury (49); biological response to stress (51, 64), and regulation of zinc and copper during development (3, 5, 52, 59). The existence of numerous independent induction responses to different agents regulating MT gene expression suggests a complex role for the regulatory recognition sequences in these genes (16, 22, 30, 43). The sequencing of members of a family of MT genes allowed us both to explore in detail the evolutionary relationship among these genes and to compare the regions of sequence involved in regulating MT gene expression.

MATERIALS AND METHODS

Screening of Charon 4a library, DNA sequence analysis, and computer analysis of sequence data. The screening of a rat liver Charon 4a genomic library for MT-1 sequences and the restriction maps for the inserts in p34 (MT-1 structural gene), p21 (MT-1ψa), p27 (MT-1ψb), and p5 (MT-1ψc) have already been reported (2). The sequence analysis of these clones was done by a combination of base-specific chemical cleavage of 5′ end-labeled restriction fragments from p34, p27, and p21 (42) and the dideoxy chain terminator method (60) after

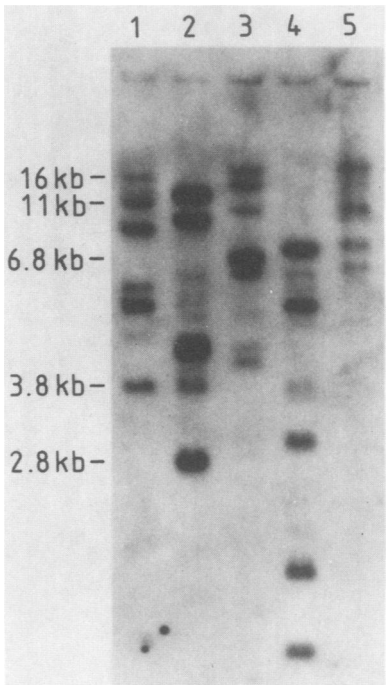


FIG. 1. Southern blot analysis of rat genomic DNA probed with a rat MT-1 cDNA sequence. A 10-μg amount of high-molecular-weight DNA isolated from H4IIE cells was digested with *Bst*EII (lane 1), *Hin*dIII (lane 2), *Eco*RI (lane 3), *Bgl*II (lane 4), or *Bam*HI (lane 5), subjected to electrophoresis through a 0.7% agarose gel, transferred to nitrocellulose, hybridized to nick-translated rat MT-1 plasmid (p2A10), and autoradiographed. Size markers were determined from *Eco*RI and *Hin*dIII digestions of lambda phage DNA.

* Corresponding author.

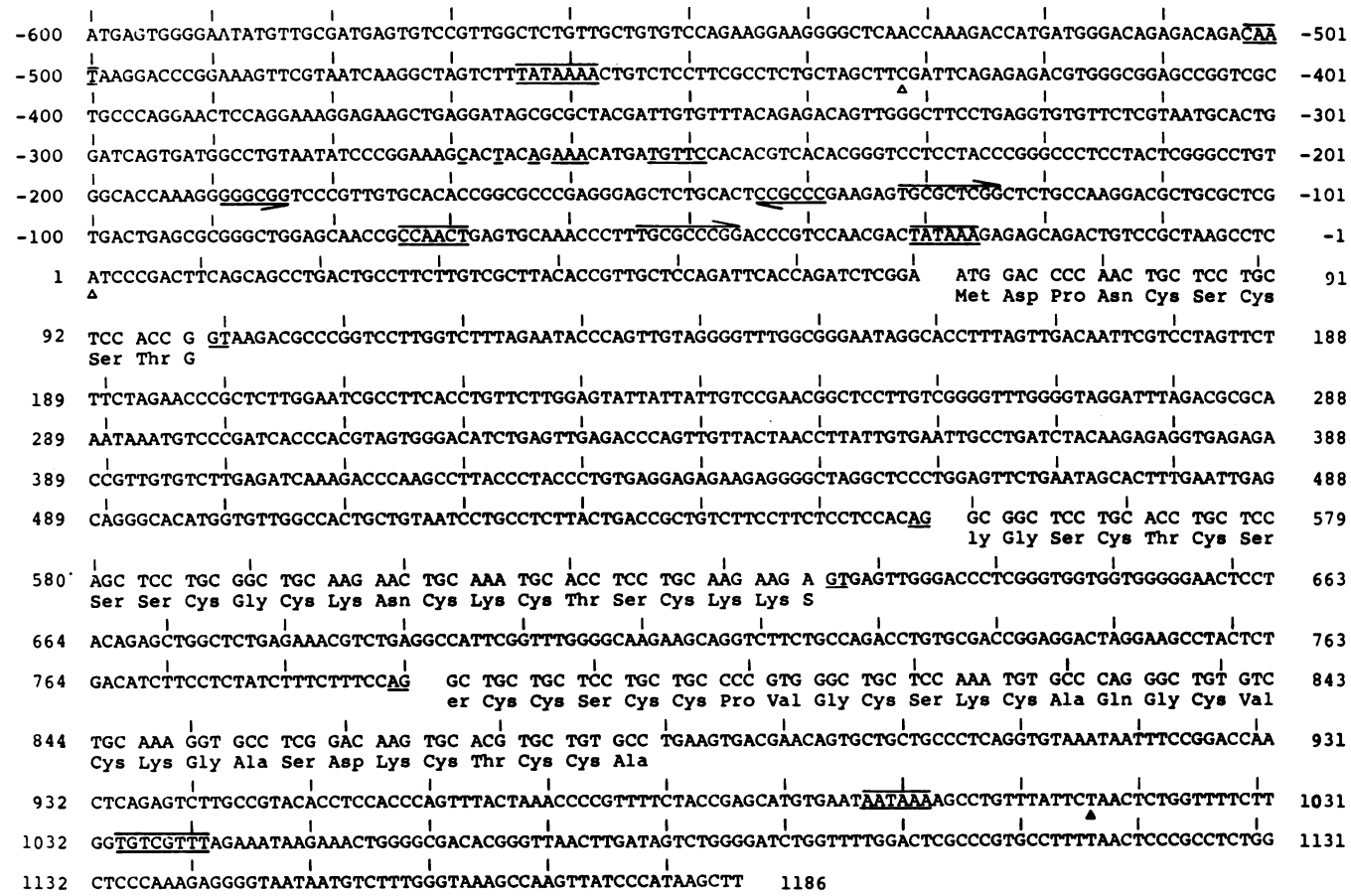


FIG. 2. Sequence of rat MT-1 structural gene and flanking sequences. This sequence includes the complete primary transcription unit, 0.6 kb of 5′- and 0.2 kb of 3′-flanking sequences. The *Eco*RI-*Hin*dIII subclone of plasmid p34 was sequenced by base-specific chemical cleavage (42) and by the dideoxy chain terminator method (60), using deletion mutants in bacteriophage M13 produced with BAL31 endonuclease (54). Data was assembled and displayed with the computer programs of Staden (66). The sequence is numbered from the beginning of exon 1. Boxed regions indicate promoter "CCAACT" and "TATA" sequences (7, 10, 14), and sequences involved in processing at the polyadenylation site (45, 55). Overlined arrows indicate metal regulatory sequences (12, 26, 34, 35, 38, 57, 61, 67). Underlined arrows indicate "GC" box inverted repeats (11, 44). Underlined nucleotides in the promoter region are homologous to the glucocorticoid binding site consensus (35).

subcloning into replicative forms of bacteriophage M13mp8 or M13mp9 or both and preparation of successive deletion mutants with BAL31 exonuclease (54). Sequence data was analyzed by using the computer programs of Staden (66) and Kanehisa (33). Each nucleotide was sequenced a minimum of two times and an average of five to six times. Complete sequence data for the 3.6-kilobase (kb) MT-1 *Eco*RI-*Hin*dIII fragment, the 2.3-kb MT-1ψa *Eco*RI-*Hin*dIII fragment, the 2.7-kb MT-1ψb *Hin*dIII fragment, and 1.1 kb at the 5′ end of the MT-1ψc *Eco*RI fragment are available on the GenBank sequence library.

Cell lines, preparation of DNA and mRNA, and DNA and RNA blot analysis. H4IIE cells (a rat hepatoma cell line [53]) were grown in modified Eagle medium. DNA was isolated from cells by the method of Gross-Bellard et al. (28). Cytoplasmic polyA$^+$ mRNA was isolated as previously described (3) from rat livers of control, dexamethasone-treated, and $CdCl_2$-treated animals and from fetal rat liver. RNA blot analysis was carried out with 1.5% agarose gels with glyoxylated mRNA by the method of Thomas (68). For genomic analyses, H4IIE DNA (10 μg) was digested with a threefold excess of restriction endonuclease, and the products resolved on a 0.7% agarose gel. After transfer to nitrocellulose (65), the blots were prehybridized and hybridized with about 2×10^7 cpm of nick-translated probe (58), hybridized overnight at 42°C (70), and washed to a final stringency of 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 50°C. 293 cells, which are derived from human embryonic kidney cells transformed by DNA from adenovirus type 5 (27), were grown in Dulbecco modified eagle medium containing 10% fetal calf serum.

S1 protection mapping. The DNA probe used for mapping the MT-1 gene trancription start site was prepared from an *Mbo*II-*Bgl*II fragment of p34 (MT-1) (−127 to +65 base pairs [bp]) inserted into the *Sma*I-*Bam*HI site of pUC13. This plasmid, pUC13MT, was digested with *Sau*3A, terminal phosphates were removed with bacterial alkaline phosphatase (Bethesda Research Laboratories), and the 5′ ends were labeled with [γ-^{32}P]ATP (ICN) and T4 polynucleotide kinase (Bethesda Research Laboratories) to a specific activity of greater than 5×10^6 cpm/pmol of 5′ ends. The DNA was then digested with *Sst*I, and the appropriate restriction fragment was isolated from a 5% polyacrylamide gel. Probes used to map the promoter site 5′ of the MT-1 gene and to detect MT-1ψb transcripts, were similarly prepared and are further described in the figure legends.

End-labeled probe and RNA were subjected to digestion by S1 nuclease (Pharmacia P-L Biochemicals) by the method

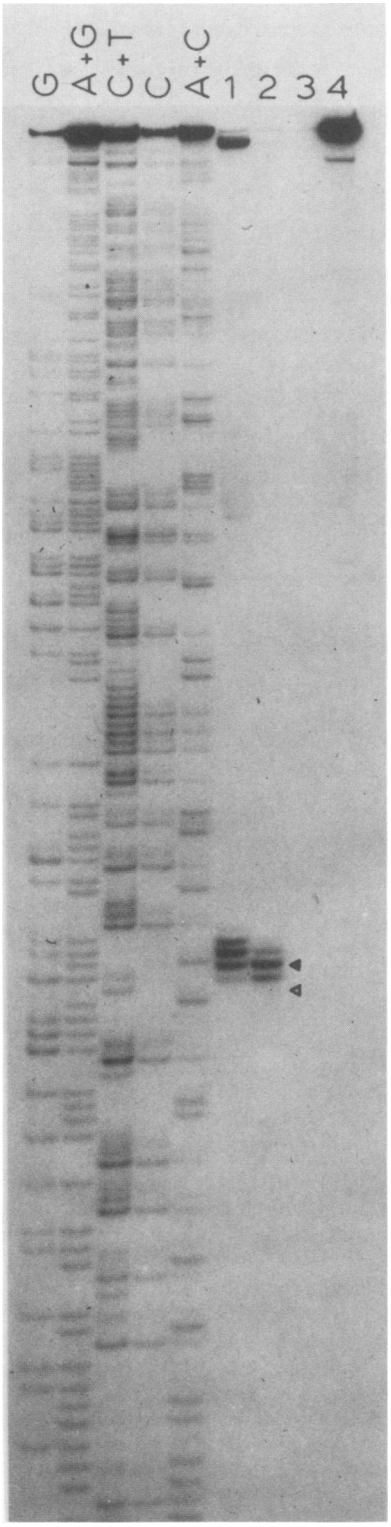


FIG. 3. S1 analysis to determine the site of transcription initiation in the rat MT-1 gene. A gene fragment from *Mbo*II to *Bgl*II (positions −127 to +65 in Fig. 2) was 5′ end labeled at the *Bgl*II position and hybridized to 3 μg of hepatic poly(A)-selected mRNA from $CdCl_2$-treated rats. The S1-resistant fragments were analyzed in a 30-by-40-by-0.025-cm 6% polyacrylamide gel containing 8.3 M urea. The probe was subjected to base-specific chemical cleavages (42) and run in parallel as size markers (first five lanes). The sequence is from the DNA strand complementary to mRNA. Lanes: 1, 100 U of S1 nuclease per ml + mRNA; 2, 500 U of S1 nuclease + mRNA; 3, 500 U of S1 nuclease per ml + 3 μg of tRNA; 4, no S1 nuclease or RNA added. The open triangle corresponds to nucleotide position +1 bp in Fig. 2. The solid triangle corresponds to the predominant site of S1 nuclease digestion, at nucleotide −2 bp, in Fig. 2.

of Berk and Sharp (8). S1-digested DNA was analyzed in polyacrylamide gels containing 8.3 M urea. Gels were fixed and dried as described by Garoff and Ansorge (24). Nuclease-resistant bands were detected by autoradiography with Kodak XRP film after overnight exposure at room temperature.

Transfection of 293 cells. The transfection protocol was adapted from the method of Van der Eb and Graham (69). Briefly, a calcium phosphate precipitate was prepared in HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid)-buffered saline (pH 7.05)–0.125 M $CaCl_2$ at a final concentration of 10 μg of DNA per ml as described by Van der Eb and Graham. A 2-ml amount of this precipitate (20 μg of DNA) was added to each 100-mm culture dish after removal of the growth medium. After 20 min, 10 ml of growth medium was added and the cells were incubated for 3.5 h. The growth medium was replaced with 15% glycerol in HEPES-buffered saline, and 30 to 60 s later this osmotic shock solution was replaced with 10 ml of phosphate-buffered saline. The phosphate-buffered saline was replaced with fresh growth medium, and the cells were incubated for 44 h.

Chloramphenicol acetyltransferase assay of cell extracts. We assayed cell extracts for conversion of [^{14}C] chloramphenicol to the acetylated form by the method of Mercola et al. (46). After separation of unreacted [^{14}C]chloramphenicol from acetylated product by thin-layer chromatography, the plates were dried, exposed to X-ray film overnight, and the conversion products were cut out and counted in Econofluor (New England Nuclear). Assays were done on three culture plates for each time point, and the average incorporation at each time point was reported.

Preparation of synthetic oligonucleotide probe. The 19-base oligonucleotide CTTCTTGCAAGGAGGTGCA was prepared at the Fermentor Laboratory at the University of California Los Angeles. A 100-pmol amount of this DNA was 5′ end labeled with [γ-^{32}P]ATP (ICN) and T4 polynucleotide kinase (Bethesda Research Laboratories). A 50-pmol amount was hybridized to a Northern blot of tissue and riboprobe RNAs, which were transferred to a Zeta Probe filter (Bio-Rad Laboratories) in 20× SSC, baked for 2 h at 80°C, prehybridized in 10× Denhardt–0.1% sodium dodecyl sulfate for 1 h at 60°C, and washed in 2× SSC. Hybridization was done in 5× Denhardt–5× SSPE–0.1% sodium dodecyl sulfate at 48°C for 3 h. The blots were washed three times in 6× SSC for 15 min at room temperature. The stringent washes were done 2× for 1.5 min at 58°C and again at 65°C in 6× SSC. After each stringent wash, an exposure was made to X-ray film.

Preparation of riboprobe RNAs. Riboprobes vectors (pSP64, from Promega-Biotech) containing either the complete MT-1 cDNA coding sequence or the *Bgl*II-*Bgl*I fragment of coding sequence from MT-1ψb were transcribed in a reaction mix containing 10 μg of vector in 52.6 μl of H_2O, 20 μl of 5× transcription buffer (200 mM Tris [pH 7.5], 30 mM $MgCl_2$, 10 mM spermidine), 1 μl of 1 M dithiothreitol (freshly prepared), 3.4 μl of RNasin (30 U/μl, Promega-Biotech), 20

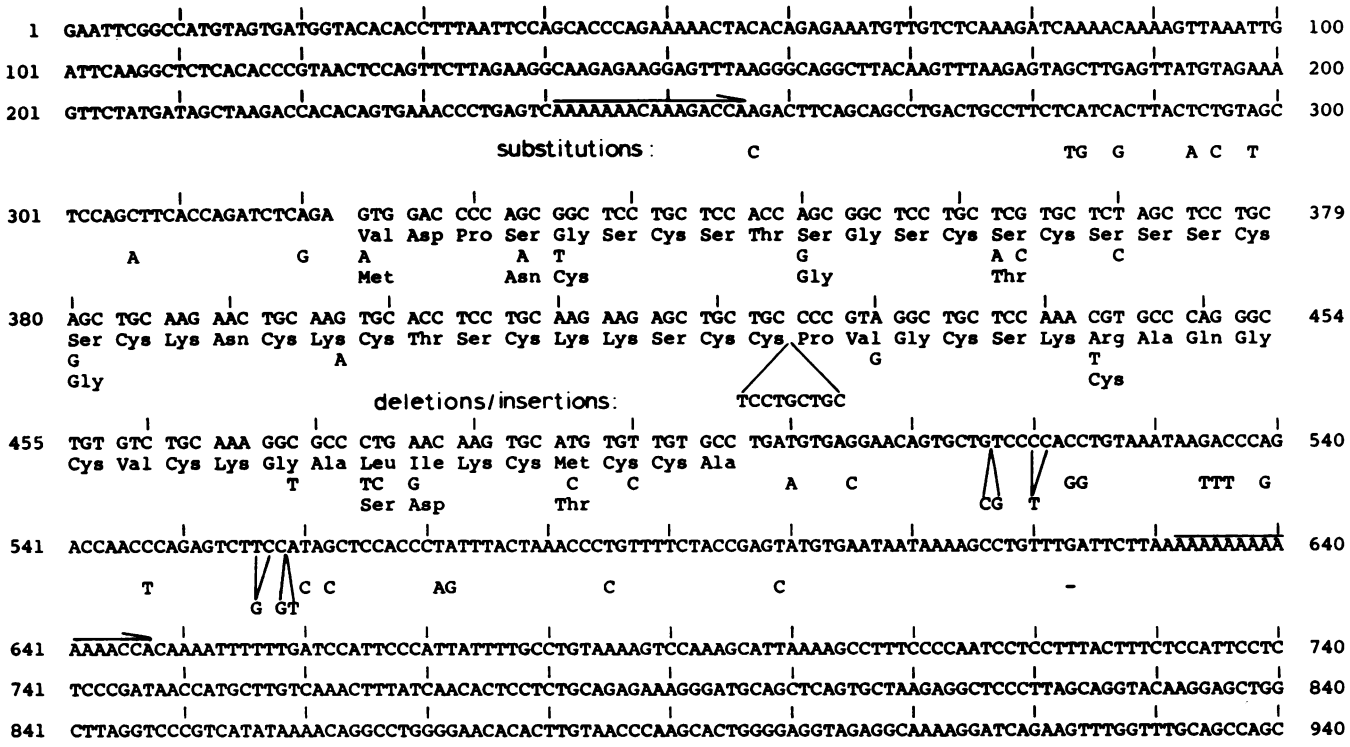
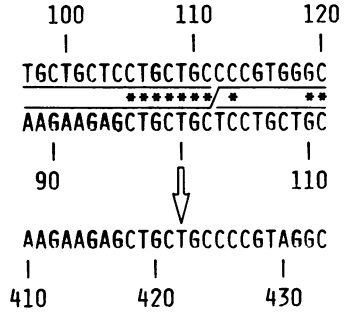
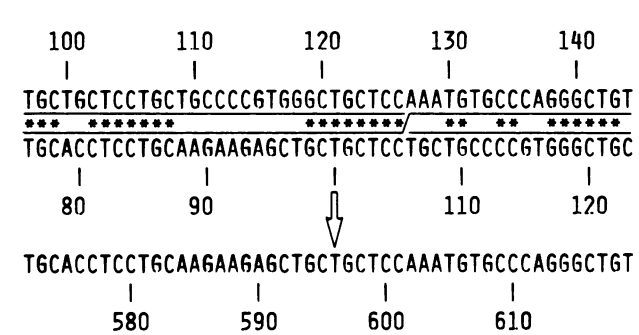


FIG. 4. Sequence of rat pseudogene MT-1ψa and flanking sequences. The *Eco*RI-*Hin*dIII insert of plasmid p21 was sequenced by the dideoxy chain terminator method described in the legend to Fig. 2. The predicted amino acid sequence is shown underneath the DNA sequence in the coding region of the gene. Underneath these are the nucleotide discrepancies with the MT-1 structural gene. Insertions relative to the structural gene are shown as a dash under the inserted base and deletions are indicated at the site of deletion, with the deleted bases underneath. Amino acid discrepancies with the MT-1 structural gene are shown under the predicted pseudogene amino acid sequence. The flanking direct repeat sequences are overlined.





μl of rNTPs (25 mM each of ATP, CTP, GTP, and UTP), and 3 μl of SP-6 polymerase (15 U/μl, Promega-Biotech). This reaction mix was prepared at room temperature and then incubated at 40°C for 1 h. The incorporation of ribonucleotides into RNA was measured by including [^{3}H]UTP (New England Nuclear) in the reaction mix. A 10-μg amount of vector gave 18 μg of 140 nucleotide MT-1ψb RNA and 39 μg of 235 nucleotide MT-1 RNA.

Preparation of metallothionein-chloramphenicol acetyltransferase fusion genes. The construct pMT(p+)cat consists of the *Mbo*II-*Bgl*II fragment (from positions −125 to 65 [see Fig. 2]) from the MT-1 structural gene cloned into pUC13 (*Sma*I-*Bam*HI) to generate *Eco*RI and *Hin*dIII ends, and this *Eco*RI-*Hin*dIII fragment, together with a CAT gene *Hin*dIII-

FIG. 5. Possible mechanism of deletion of direct repeat sequence segments in pseudogenes MT-1ψa and MT-1ψc. The two upper sequences in each part of this figure depict a region of the MT-1 cDNA (numbered from the transcription initiation site of the MT-1 gene), the lower of the two sequences being the upper sequence shifted 9 or 21 nucleotides (MT-1ψa and c, respectively) to the right. Asterisks (*) indicate the alignment of identical nucleotides; a row of these indicates a region of direct repeat in the cDNA. The third row indicates the predicted sequence which would occur after a deletion caused by a slipped mispairing mechanism (1). The numbering corresponds to the actual sequence found in each pseudogene. (Note: nucleotide 430 in MT-1ψa has independently mutated from a G to an A, Fig. 4). The resulting deletions involve the codons SerCysCys (MT-1ψa) or CysCysProValGlyCysSer (MT-1ψc), where the underlined amino acids are common to both sequences.

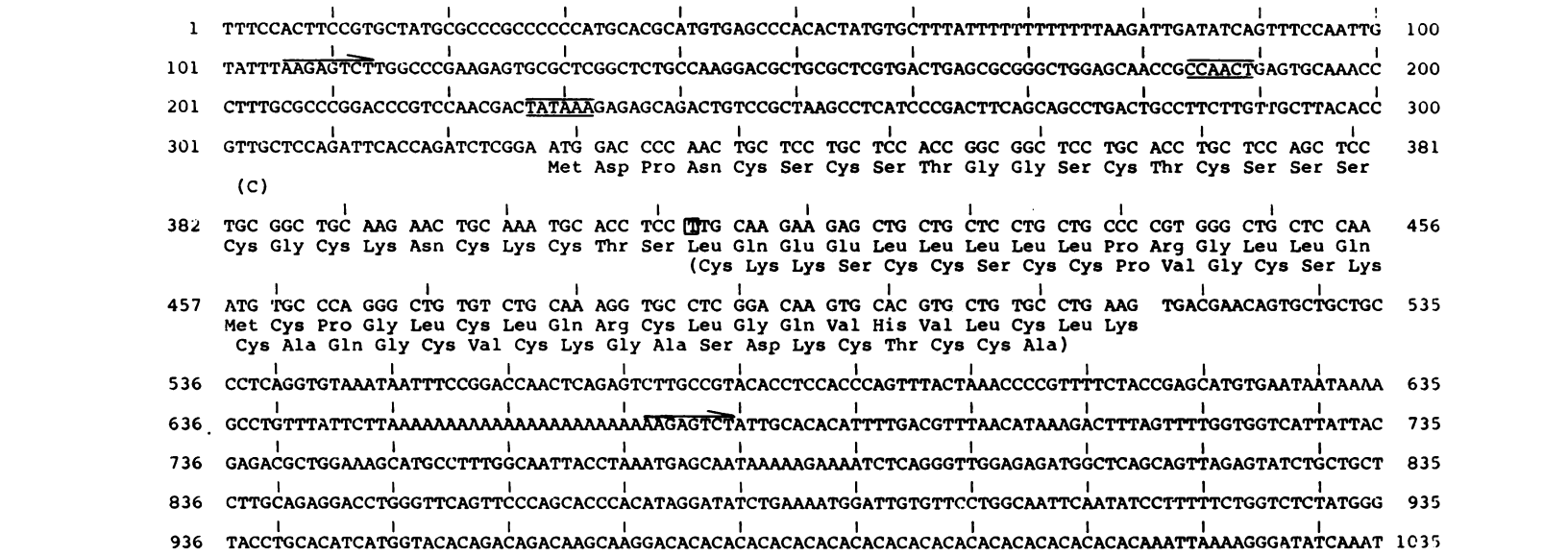


FIG. 6. Sequence of rat pseudogene MT-1ψb and flanking sequences. The *Hin*dIII insert of plasmid p27 was sequenced by a combination of base-specific chemical cleavage (42) and the dideoxy chain terminator method (60), using deletion mutants in bacteriophage M13 produced with BAL31 endonuclease (54). The predicted amino acid sequence in the coding region of the gene is shown underneath the DNA sequence. The single nucleotide substitution within the MT-1 gene is shown below the pseudogene sequence. The site of the T insertion is boxed. The MT-1 amino acid sequence beyond this site is shown in parentheses under the amino acid sequence of the pseudogene. The flanking direct repeats are overlined. The "CCAACT" and "TATA" consensus sequences are open boxed.

*Bam*HI fragment (isolated from $pA_{10}cat_2$), cloned into pUC13 (*Eco*RI-*Bam*HI). The construct pMT(ψb)cat consists of a 1.1-kb *Hin*dIII-*Bgl*II fragment (the *Bgl*II site [3′] lies at position 322 in Fig. 6; the *Hin*dIII end [5′] was polished with T4 polymerase) cloned into pUC13 (*Eco*RI-*Bam*HI).

RESULTS

A multigene family of metallothionein-related sequences in rats. We probed a Southern blot of rat DNA digested with the restriction enzymes *Bst*EII, *Eco*RI, *Hin*dIII, *Bam*HI, or *Bgl*II (Fig. 1) with the complete rat MT-1 cDNA sequence (2) and observed at least six bands homologous to this probe with each restriction digest. By using sequences unique for the MT-1 structural gene and for three MT-1 pseudogenes (see sequence analyses below) we identified bands corresponding to the MT-1 structural gene (3.8-kb *Bst*EII, 6.8-kb *Eco*RI, and 4.6-kb *Hin*dIII fragments), and three pseudogenes: MT-1ψa (16-kb *Bst*EII fragment); MT-1ψb (11-kb *Bst*EII and 2.8-kb *Hin*dIII fragments); and MT-1ψc (5.6-kb *Bst*EII fragment) (unpublished observations). The human MT-2 genes also show a similar complexity on Southern blot analysis (38). A Southern blot of mouse genomic DNA has also been reported (6). Only the mouse MT-1 genomic sequence hybridized when probed with the mouse MT-1 gene. One of us (31) has repeated this experiment at the same stringency used in Fig. 1, using DNA from mouse 3T3 cells and a murine cDNA probe, and confirms the absence of any MT-1 related genomic sequences other than the MT-1 structural gene.

Sequence analysis of the MT-1 structural gene. Isolation of MT-1 sequences homologous to an MT-1 cDNA probe from a Charon 4a rat *Hae*III genomic library has been described previously (2). The sequence of a 3.6-kb *Eco*RI-*Hin*dIII fragment containing this gene subcloned into pBR322 (designated as subclone p34) has been analyzed. In Fig. 2 we present the 1.8 kb at the 3′ end of this sequence.

Comparison of the MT-1 genomic sequence with the MT-1 cDNA sequence shows the presence of two intervening sequences and exon-intron junctions which obey the GT-AG rule (9). The 3′ end of each intron contains a short tract of pyrimidines also observed in the introns of other genes (40). The 5′ end conforms to the GT(A,G)AGT consensus (48). The transcription initiation site was determined by S1 nuclease mapping (8, 71), using a ^{32}P-labeled probe and rat mRNA from $CdCl_2$-treated animals. The sizes of the resistant fragments were compared to a sequence ladder of the MT-1 gene. We detected a series of bands, suggesting that the S1 nuclease is encountering steric hindrance at the cap site of the MT-1 mRNA. This was confirmed when a comparison was made between digestion at S1 nuclease concentrations of 100 and 500 U/ml (Fig. 3, lanes 1 and 2, respectively). When we examined this region of the MT-1 sequence (Fig. 2), we found two possible locations for the cap site. The predominant S1-protected band we indicated with a solid triangle (Fig. 3); the furthest extent of S1-nuclease digestion we have indicated with an open triangle (Fig. 2 and 3). This latter position would agree with the observation of a Py-C-A-Py consensus sequence found for eucaryotic mRNAs (14) at the cap site. A similar result was obtained with the mouse MT-1 gene (26). The corresponding site of predominant S1 nuclease digestion for mouse MT-1 mRNA was found to be one nucleotide upstream from the site of predominant S1 digestion found with rat MT-1. We numbered the rat MT-1 gene assuming the cap site maintains agreement with the Py-C-A-Py rule (open triangle at position +1 in Fig. 2 and 3).

A comparison of the rat MT-1 cDNA sequence (2) with the genomic sequence (Fig. 2) shows a single replacement in the 5′ noncoding region of exon 1 of the MT-1 gene in which a T substitutes at position 18 for a C in the cDNA sequence. We assume this is a sequence polymorphism between the cDNA and genomic DNA sources. A comparison between the rat MT-1 cDNA and gene sequences reveals that the site of polyadenylation lies after position 1016, 14 to 16 bp down-

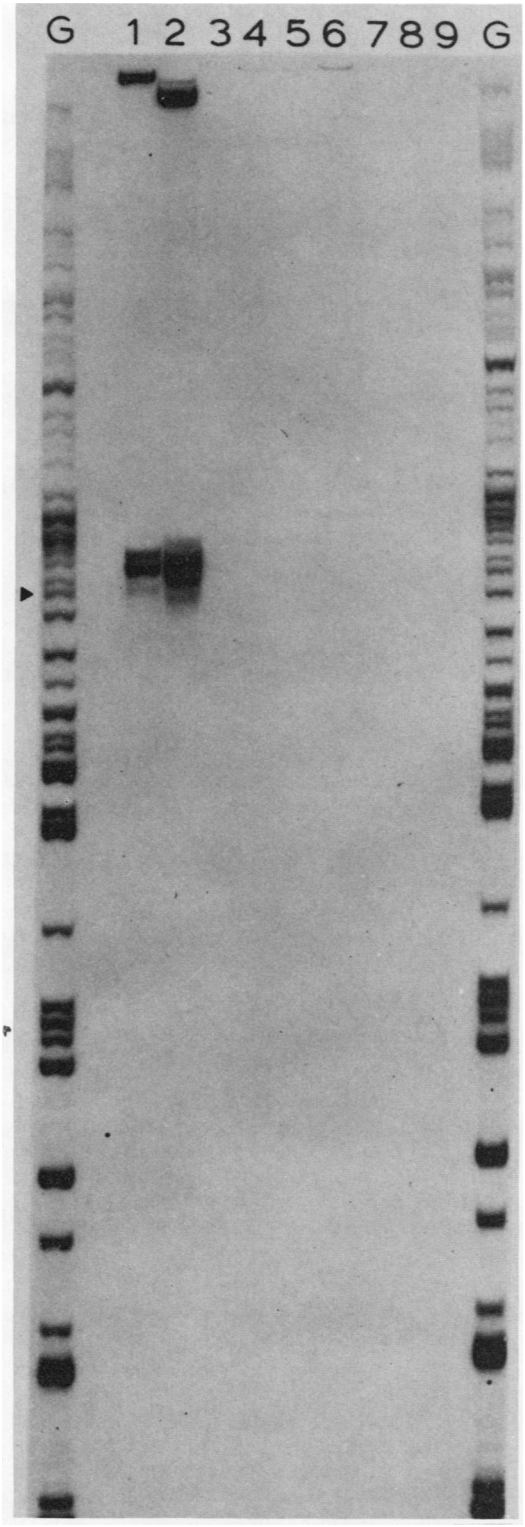


FIG. 7. S1 analysis to determine the presence of a promoter 5′ of the MT-1 structural gene sequence. A gene fragment from *Dra*I to *Sma*I (positions −679 to −225) in the MT-1 structural gene was 5′ end labeled at the *Sma*I position and hybridized to various RNAs. Digestions were done at 500 U of S1 nuclease per ml. The S1-resistant fragments were analyzed as described in the legend to Fig. 4. Flanking the lanes containing protected fragments are G ladders from the Maxam and Gilbert G-specific chemical cleavage of the end-labeled probe. In lanes 1 and 2 are total cellular RNAs from 293 cells transfected with p34 (MT-1) DNA at 20 μg/100-mm plate. In lane 2, the 293 cells were treated for 6 h with 5 μM $CdCl_2$ before harvest. Lane 6, Total cytosol RNA from H4IIE cells. Lanes 3 to 5 and lane 7, poly(A)$^+$ mRNAs isolated from (lane 3) control rat liver, (lane 4) rat liver from animals injected with 10 μmol of $CdCl_2$ per kg of body weight 5 h before sacrifice, (lane 5) rat liver from animals injected with 68 μmol of hydrocortisone sodium succinate + 50 μmol of hydrocortisone per kg of body weight 5 h before sacrifice, (lane 7) rat liver from 17-day gestation fetuses. In protection experiments 0.4 pmol of probe was hybridized to 50 μg of total cellular RNA (lanes 1 and 2), 30 μg of total cytosol RNA (lane 6), and 5 μg of poly(A)$^+$ mRNA (lanes 3 to 5 and 7). The triangle corresponds to position −432 bp in Fig. 2.

stream from the consensus sequence AATAAA (solid triangle in Fig. 2).

Sequence analysis of a processed MT-1 pseudogene, MT-1ψa, resulting from reverse transcription of MT-mRNA. We have previously described the isolation, from the Charon 4a rat genomic library, of a gene which shows some homology to the rat MT-1 sequence (2). The sequence of a 2.3-kb *Eco*RI-*Hin*dIII fragment subcloned into pBR322 (designated as subclone p21) containing this gene has been analyzed. A portion of that sequence, from the *Eco*RI site to a point 900 bp 3′ of this site, is shown in Fig. 4. Several features of this sequence identify it as a processed MT-1 pseudogene and suggest its mechanism of formation. The pseudogene of p21, which we called MT-1ψa, demonstrates the absence of intervening sequences, the presence of a polyadenylation site at a position 17 bp downstream from the AATAAA consensus sequence (position 630), and a direct repeat of 16 bp involving this polyadenylation site, within 4 bp of the transcription initiation site (position 257). These features are diagnostic of a processed pseudogene in which the MT-1 mRNA transcript was reverse transcribed to produce an MT-1 cDNA, which was subsequently inserted into the rat genome.

Homology between MT-1ψa and the MT-1 exon regions extends for 372 nucleotides, with 331 matches, 40 replacements, 15 deletions, and 2 insertions in MT-1ψa. There is 86% overall homology for the 371 bp compared between the MT-1 structural gene and MT-1ψa. A deletion of 9 bp occurs between positions 424 and 425. None of these changes affects the reading frame of the coding region in this gene, nor do any changes introduce a premature stop codon. Since evidence has been accumulated which shows a rapid and constant rate of divergence of pseudogene sequences of 400 PAMs (R. F. Doolittle, personal communication; 1 PAM = 1 accepted point mutation per 100 residues per 100 million years) calculated from amino acid data, we were able to compute the approximate time of divergence of this pseudogene from its parent MT-1 gene. Translating the coding sequence into amino acids (see Fig. 4), and comparing this sequence to the rat MT-1 amino acid sequence, we found 10 substitutions and 1 deletion, for a total of 11 mutational events. For a protein which originally had 61 amino acids, we estimated the time of divergence as 4.5 million years ago. This is about half the age of divergence of mice and rats (56) and is probably a conservative estimate since we have ignored independent mutational events at a single site or superimposition (15). An alternative calculation can be made from the DNA sequence data in the coding region. Adding up the mutations at all three codon positions, plus the deletion, we obtained 18 mutations. The base substitution rate for pseudogenes is approximately twice the rate for

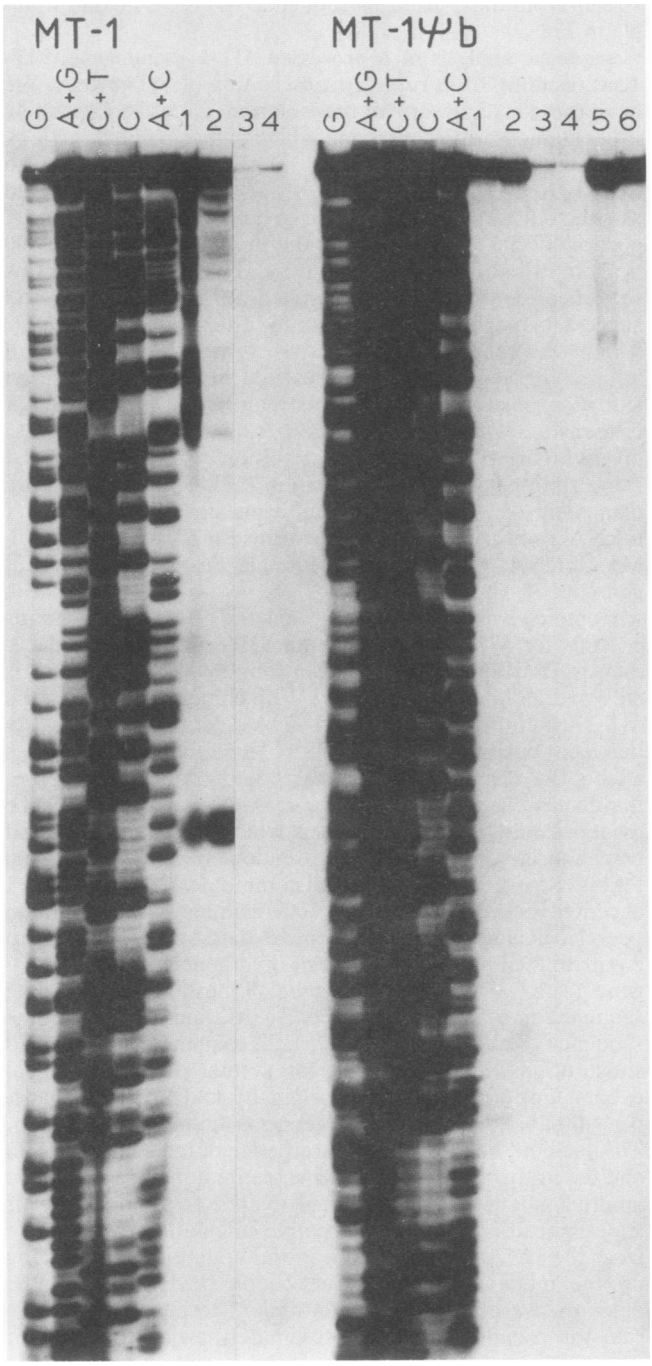


FIG. 8. S1 analysis to detect MT-1 gene and pseudogene MT-1ψb transcripts in transfected cells. Gene fragments from *Mbo*II to *Bgl*II in p34 (positions −125 to +65 bp in Fig. 2) and in p27 (133 to 322 bp in Fig. 5) were 5′ end labeled at the *Bgl*II position, and 0.4 pmol was hybridized to 50 μg of RNAs isolated from 293 cells transfected with either p34 (MT-1) or p27 (MT-1ψb) plasmid DNAs. Digestions were done at 500 U of S1 nuclease per ml. The S1-resistant fragments were analyzed as described in the legend to Fig. 3. On the left are protected fragments from hybridizations involving RNA from 293 cells transfected with p34 (MT-1) DNA; on the right are similar experiments done using p27 (MT-1ψb) as transfecting DNA. The sequencing reactions are shown in lanes to the left of the S1-protected fragments. Cells were transfected with 20 μg of plasmid DNA per 100-mm plate and left untreated or treated for 6 h with inducer before RNA isolation. MT-1 gene transfection (lanes): 1, p34 control; 2, p34 + 5 μM $CdCl_2$; 3 to 4, tRNA control. Pseudogene MT-1ψb transfection (lanes): 1 and 5, p27 control; 2 and 6, p27 + 5 μM $CdCl_2$; 3-4, tRNA control. In lanes 1 and 2, cytosol RNA was hybridized; in lanes 5 and 6, total cellular RNA was hybridized.

synonymous changes in functional genes (41, 47). Thus, two-thirds of the third codon mutations in MT-1ψa are due to changes in the pseudogene after integration, and one-third are due to evolutionary drift in the functional gene. This reduces the total mutational differences to 16, and gives a rate of divergence of pseudogene MT-1ψa from the parent gene of 8.7%, which corresponds to an evolutionary age of 6.9 million years. Either calculation suggests that MT-1ψa probably appeared after the divergence between rats and mice, consistent with data from Southern blots of DNA from these two species.

Similar deletions characterize two processed pseudogenes, MT1ψa and MT1ψc. We previously reported the cloning from the Charon 4a rat genomic library of a 2-kb *Eco*RI fragment which shows weak homology to the rat MT-1 sequence (2). We sequenced part of the 2-kb *Eco*RI fragment subcloned into pBR322 (designated as subclone p5) and identified a portion of this sequence as a second MT-1 retrogene, MT-1ψc. The 5′ end of this pseudogene (sequence not shown; available on GenBank) exhibits homology to the rat MT-1 sequence beginning with the transcription initiation site. Unlike MT-1ψa, MT-1ψc lacks 3′ sequences homologous to the MT-1 gene past the sixth nucleotide 3′ of the stop codon. A total of six substitutions and one deletion in the coding region of MT-1ψc (of which four are in the first and second codon positions and three are in the third) yielded a rate of divergence (see above) of MT-1ψc from the parent gene of 3.3%, corresponding to an evolutionary age of 2.6 million years.

Pseudogenes MT-1ψa and MT-1ψc both showed deletions in the same region of coding sequence (Fig. 5). When regions of the MT-1 cDNA covering the deleted sequence for each pseudogene are compared with sequences either 9 nucleotides (MT-1ψa) or 21 nucleotides (MT-1ψc) 3′ to the initial sequence occurrence, a region of direct repeat sequence is found which terminates at the beginning of the homologous region deleted in each pseudogene (Fig. 5). This observation suggests that the presence of short repeated sequences in the rat MT-1 coding region results in an increased likelihood of deletion formation, once the selective pressure to maintain that coding sequence has been removed by pseudogene formation. The importance of short repeated sequences in specifying the formation of spontaneous deletions has been observed previously (1). These deletions in MT-1ψa and MT-1ψc are mechanistically plausible by the slipped mispairing scheme for spontaneous deletion formation proposed by Albertini et al. (1).

Sequence analysis of an MT-1 pseudogene, MT-1ψb, which includes the MT-1 structural gene promoter sequence. We previously reported a pseudogene, MT-1ψb, isolated from the rat genomic library (2). The sequence of a portion of a 2.8-kb *Hin*dIII fragment, subcloned into pBR322 (designated subclone p27), is presented in Fig. 6. This sequence shows the same features of a processed pseudogene seen with MT-1ψa, namely, (i) a lack of intervening sequences, (ii) a polyadenylation tract 3′ to the transcription terminator, and (iii) direct repeats flanking the region homologous to the parent gene. In addition, both pseudogenes MT-1ψa and MT-1ψb show the addition of an extra T residue which is not

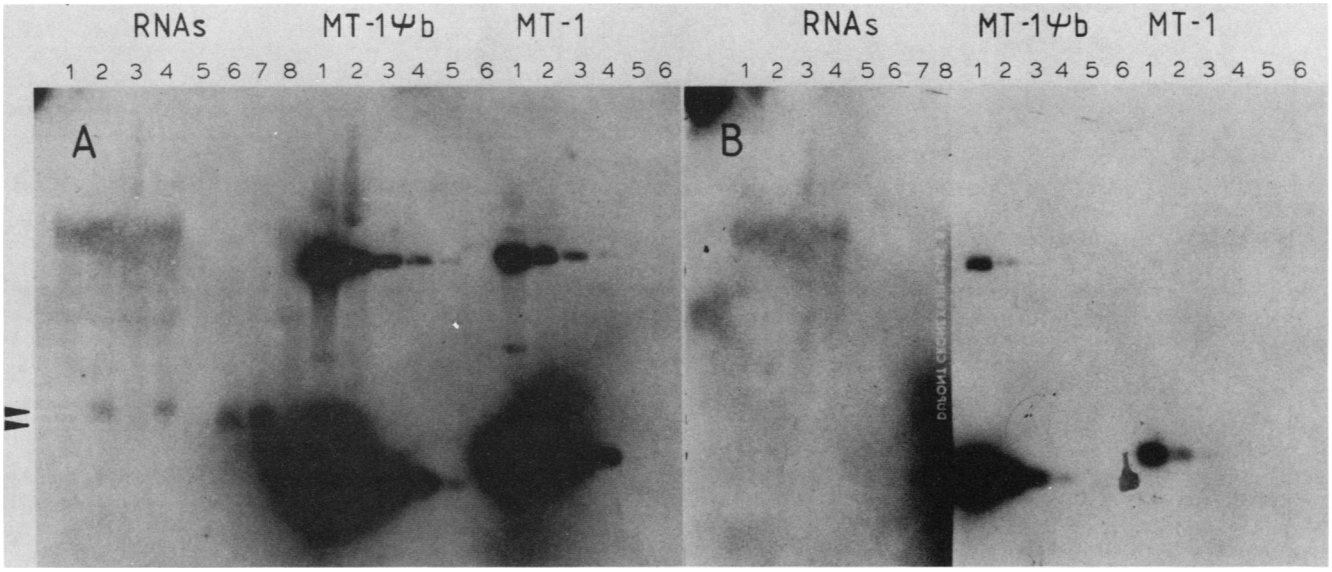


FIG. 9. Northern blot analysis to probe for pseudogene MT-1ψb transcripts in liver tissue or transfected cells. Varous RNAs were electrophoresed in a 1.5% agarose gel, transferred to nitrocellulose, and probed with 50 pmol of the end-labeled sequence CTTCTTGCAAG-GAGGTGCA at a temperature of 58°C (Fig. 9A) or 65°C (Fig. 9B). The RNAs (lanes 1 to 8 on the far left of each panel) were isolated from: 293 cells transfected with p27 (MT-1ψb) DNA at 20 μg/100-mm dish (lanes 1 and 2, 20 μg of cytosol RNA; lanes 3 and 4, 20 μg of total cellular RNA); from rat liver tissue (lanes 5 to 7); from H4IIE rat hepatoma cells (lane 8, 20 μg of cytosol RNA). The various treatments and sources of RNA are (lanes): 1 and 3, p27 control; 2 and 4, p27 + 5 μM $CdCl_2$ (6-h exposure); 5, control rat liver; 6, rat liver isolated 5 h after injection of 10 μmol of $CdCl_2$ per kg of body weight; 7, rat liver isolated 5 h after injection of 68 μmol of hydrocortisone sodium succinate + 50 μmol of hydrocortisone per kg of body weight; 8, H4IIE cells + 3 μM $CdCl_2$. The lanes in the center and on the right in each panel are RNAs synthesized from riboprobe vectors containing MT-1 or MT-1ψb sequences of 235 and 140 bases, respectively (see text). The amount of these sequences applied is (lanes): 1, 10 pmol; 2, 2.5 pmol; 3, 0.6 pmol; 4, 0.16 pmol; 5, 0.04 pmol; 6, 0.01 pmol. The lower bands in these lanes correspond to RNAs synthesized from the riboprobe vectors. The upper bands are vector DNA. The cellular RNAs on the left in Fig. 9B (after the 65°C wash) were exposed sixfold more with the use of an intensification screen. The positions of human (lanes 2 and 4) and rat (lanes 6, 7, and 8) cellular mRNAs observed as a result of hybridization with the probe at the lower stringency are indicated by the upper and lower arrows, respectively (see text).

present in either the MT-1 structural gene or the cDNA sequences at the polyadenylation site.

The sequence of MT-1ψb shows several features not normally seen in pseudogenes. There is only a single substitution, C for T, at position 291 in the 5′ noncoding region. There are no base substitutions throughout the coding region; however, there is a single insertion of a T at position 412. The T insertion produces a frameshift in the middle of the coding region, resulting in a new open reading frame encoded by this pseudogene. This frameshift also causes a new termination codon to occur, 3 bp 3′ to the normal MT-1 stop codon. A new coding sequence is therefore created, in which a molecule homologous to MT-1 in its N-terminal end but with a unique amino acid C-terminal end is predicted. The last sequence difference is the presence of two mismatches at the 5′ end and one mismatch at the 3′ end of this sequence, inside the AAGAGTCT direct repeats. The close homology between this pseudogene and the MT-1 cDNA sequences, in both coding and noncoding regions, indicates that MT-1ψb is evolutionarily very young compared with MT-1ψa or MT-1ψc.

Surprisingly, MT-1ψb continues to be homologous to the structural gene sequence for 142 bp 5′ of the transcription initiation site for the normal MT-1 structural gene (located in pseudogene MT-1ψb at position 258 in Fig. 6). This homologous stretch of sequence encodes the entire region necessary for a functional promoter capable of responding to metal ion inducers (11, 35, 43). Included in this pseudogene "promoter" are the "TATA" and "CCAACT" consensus sequences and the 9 bp direct repeat consensus sequences found in all MT promoters. We conclude that this pseudogene has the regulatory sequences required both for basal transcription and for induced transcription in response to metal ion induction. The expression of this gene may be limited by its position in the rat genome, by the loss of intervening sequences, or by the presence of a polyadenylation tract at its 3′ end in place of the termination consensus sequence found in the MT-1 gene.

DNA sequences adjacent to pseudogene MT-1ψb display several identifiable features. A poly(CA) tract 42 bp long is found starting at position 973, 293 bp 3′ from the distal direct repeat. A stretch of repeating CA residues has been found in eucaryotic DNA at the site of recombination events, such as gene conversion (62), and is predicted to cause a Z-DNA structure to occur (50). The 3′ end of the 2.8-kb *Hin*dIII fragment containing MT-1ψb is a repeated sequence in the rat genome (determined by genomic Southern blot analysis, data not shown) and shows a poly(GTT) tract 18 bp long. The region of pseudogene MT-1ψb insertion in the rat genome shows both poly(T) (position 65 to 76) and poly(A) (position 650 to 673) tracts, a feature often seen at pseudogene insertion sites (E. F. Vanin, Biochim. Biophys. Acta, in press).

Possible origin of pseudogene MT-1ψb. A *Pol*II initiation site found upstream of the MT-1 structural gene. In considering the origin of MT-1ψb, we have looked for possible promoter sites located 5′ to the normal transcription initiation site in the MT-1 structural gene. Such a site might have

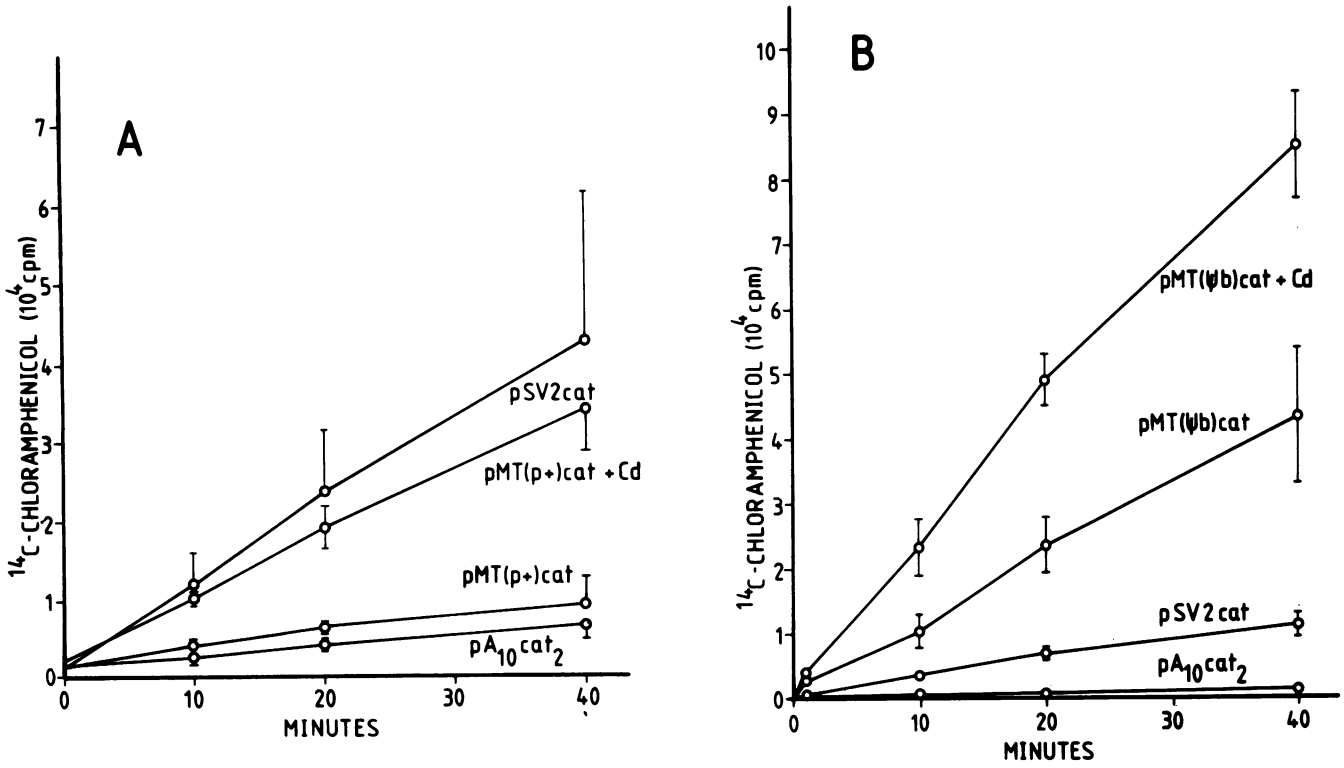


FIG. 10. Constructs which contain the MT-1 promoter sequence or MT-1ψb 5′ sequences fused to the CAT gene can be regulated by metal ion. Fusion plasmids containing the MT-1 promoter sequence, pMT(p+)cat (A), or 1.1-kb *Hin*dIII-*Bgl*II 5′ end of MT-1ψb (B) fused to the CAT gene were used to transfect 293 cells. The constructs pA$_{10}$cat$_2$ and pSV2cat were used as controls of low and high levels of CAT expression, respectively. 20 μg of DNA was transfected per 100-mm plate of 293 cells. $CdCl_2$ treatment was the same as in Fig. 8. The plot shows CAT activity as a time course of accumulation of acetylated [^{14}C]chloramphenicol, measured as ^{14}C counts per minute recovered after a thin-layer chromatographic separation.

given rise to an mRNA transcript in germline tissue which contains the MT-1 regulatory sequences and which ends at the MT-1 gene transcription termination site. An examination of the MT-1 gene sequence 5′ to the promoter region shows a TATAAAA sequence at position −464 and a CAAT sequence at position −503 (Fig. 2). We identified a transcript originating from this putative promoter sequence by S1 nuclease mapping (Fig. 6), using a probe end labeled at an *Sma*I site (−225 bp in Fig. 2) and extending in the 3′ direction to a *Dra*I site (−679 bp, sequence not shown). RNA from 293 cells (chosen for their high level of expression of exogenously added DNA [25]) transfected with the MT-1 gene (p34) showed the presence of a protected fragment. The initiation of transcription is at nucleotide −432 (Fig. 2), 32 bp 3′ from the start of the TATAAA sequence at position −464. No transcripts were detected from this region upstream of the MT-1 gene in polyA$^+$ mRNAs from control, $CdCl_2$-treated, or dexamethasone-treated rats, from total RNA from H4IIE cells, or from fetal rat liver poly(A)$^+$ mRNA. (Fetal rat liver expresses metallothioneins at elevated levels [3]). A comparison of this upstream transcript from control and $CdCl_2$-treated 293 cells, transfected with p34, showed an increased level of this RNA after metal ion treatment.

Detectable levels of MT-1ψb transcripts not found in rat liver or after transfection of the pseudogene. We performed two experiments which show that pseudogene MT-1ψb transcripts are not detectable after transfection of p27 (contains MT-1ψb) into 293 cells or in liver (i.e., where MT-1ψb is in its endogenous state). In the first of these experiments we compared the S1 nuclease mapping of transcripts from 293 cells transfected with p34 (MT-1) and p27 (MT-1ψb) plasmid DNAs (Fig. 8). In each case the probe was end labeled at the *Bgl*II site (position 65 in MT-1 and position 322 in MT-1ψb) and extended upstream to an *Mbo*II site (position −127 in MT-1 and position 133 in MT-1ψb). The transfected MT-1 gene shows induction by cadmium ion. No transcripts are seen from the transfected MT-1ψb pseudogene.

In an experiment to determine whether the MT-1ψb gene is expressed in its endogenous location in the rat genome, a Northern blot was prepared with RNAs from (i) 293 cells transfected with p27, (ii) rat liver tissue, and (iii) H4IIE cells (Fig. 9). These RNAs were from both control and $CdCl_2$-treated sources. Also on this Northern blot were two synthetic RNAs prepared from riboprobe vectors which contain coding sequence regions from pseudogene MT-1ψb (*Bgl*II-*Bgl*I, positions 318 to 438) and from the MT-1 cDNA (*Bgl*II/*Dde*I, positions −10 to 209, Fig. 2 in reference 2). This Northern blot was probed with an end-labeled 19-base oligonucleotide which overlaps the T-insertion site at nucleotide position 412 in the MT-1ψb gene: ^{32}P-CTTCTTGCAA GGAGGTGCA, and which is complementary to the mRNA sequence predicted from this pseudogene. The underlined nucleotide corresponds to the T-insertion site in the mRNA strand of the pseudogene. When this blot was washed at 58°C, MT-1 RNA was detectable at a level of less than 0.16 pmol (Fig. 9A, MT-1 lanes 4 and 5). Thus, the amount of induced MT-1 mRNA on this blot is approximately 0.1 pmol, corresponding to approximately 1.5% of total mRNA (Fig.

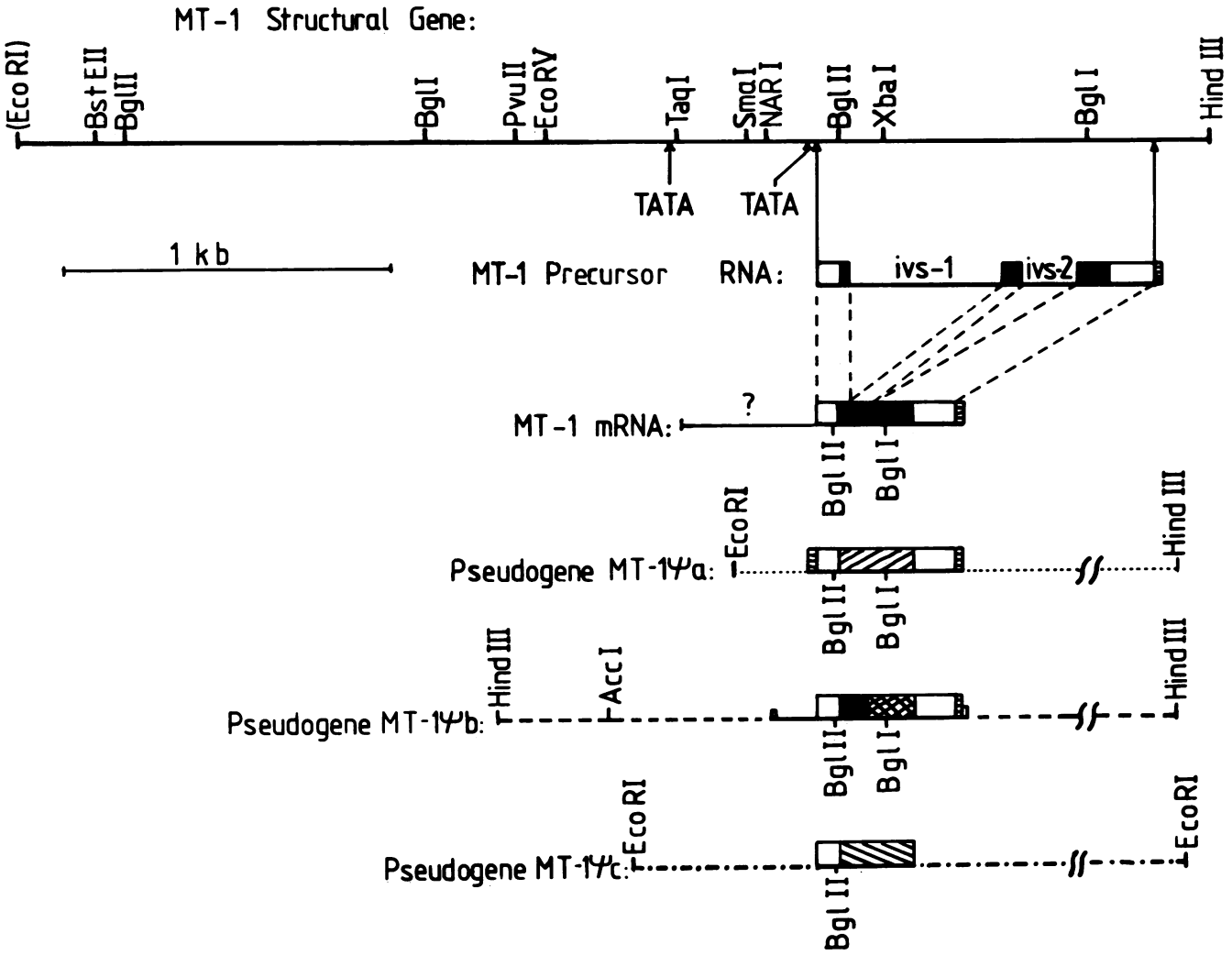


FIG. 11. Relationships between the MT-1 structural gene and pseudogenes MT-1ψa, MT-1ψb, and MT-1ψc. This figure indicates how each pseudogene may have derived from processed mRNAs. The solid bar regions are coding sequences identical to the MT-1 gene. The open bar regions are 5′ and 3′ untranslated sequences. The crosshatched region indicates that portion of pseudogene MT-1ψb coding sequence which underwent a frameshift due to a T insertion. The diagonally shaded regions are the coding sequences of pseudogene MT-1ψa or MT-1ψc, which, in each case, have undergone numerous substitutions and a deletion. The horizontally shaded regions are the locations of homopolymer A sequences. The small boxes flanking the MT-1ψb gene are the AAGAGTCT direct repeats. The sequence with a question mark, contiguous with the MT-1 mRNA transcript, indicates a hypothetical larger message derived from transcription at the upstream TATA box promoter in the MT-1 structural gene. The solid line represents genomic sequences associated with the structural MT-1 gene. The dotted line is flanking DNA at the site of insertion of the MT-1ψa sequence; the dashed line is flanking DNA at the site of insertion of the MT-1ψb gene. IVS-1 and IVS-2 are the introns of the MT-1 gene. The *Eco*RI site at the 5′ end of the MT-1 gene represents the cloning insertion site in the Charon 4a vector.

9A, RNA lanes). At 58°C the MT-1ψb RNA is detectable at a level of 0.04 pmol (Fig. 9A, MT-1ψb lane 5). When the stringency wash was increased to 65°C, the MT-1 RNA was just detectable at 0.6 pmol (Fig. 9B, MT-1 lane 3), and the MT-1ψb RNA was detectable at 0.16 pmol (Fig. 9B, MT-1ψb lane 4). From these results, we predicted that the strongest MT-1 RNA bands seen in the low stringency wash should be more than sixfold below the level of detectability with the synthetic probe at 65°C (Fig. 9B, MT-1 lane 3). The use of an intensification screen on the RNA lanes in Fig. 9B increased the sensitivity approximately sixfold. Thus, MT-1 mRNA should fall just below the level of detection in Fig. 9B (RNA lanes). The absence of any detectable hybridization in the RNA lanes at 65°C indicated that the MT-1ψb transcripts must be less than 0.16 pmol/6 = 0.03 pmol (Fig. 9B, MT-1ψb lane 4). We conclude that MT-1ψb transcripts can represent, at most, 0.1% of total mRNA.

Only 126 bp of sequence 5′ to the MT-1 gene transcription start site needed for regulation of transcription by cadmium ion. To determine whether the 142 bp of MT regulatory sequence contained in pseudogene MT-1ψb is capable of directing transcription and of regulating transcription by metal ions, we constructed a fusion gene between the rat MT-1 gene upstream sequences which are identical to those of MT-1ψb and the chloramphenicol acetyltransferase (CAT) gene (39). We called this construct pMT(p+)cat; it contains the MT regulatory sequence and the mRNA 5′ untranslated sequence, followed by the CAT gene. We transfected 293 cells with this construct, as well as with $pA_{10}cat_2$ and pSV2cat, and assayed cell extracts for CAT enzyme activity

48 h later. Plasmids $pA_{10}cat_2$ and pSV2cat contain the CAT gene under regulation by the simian virus 40 promoter, without and with the 72-bp enhancer sequences, respectively, and yield constitutively low ($pA_{10}cat_2$) and high (pSV2cat) levels of CAT activity (39). Cells transfected with the construct pMT(p+)cat showed low CAT enzyme activity in the absence of inducer, and high CAT enzyme activity after a 6-h exposure to 5 μM $CdCl_2$ (Fig. 10A). Thus, the presence of only 126 bp of sequence upstream of the MT-1 gene transcription start site is sufficient to specify a promoter which responds to metal ion as an inducer.

Regulatory region of pseudogene MT-1ψb transcriptionally active when fused to the CAT gene. One explanation for the lack of detectable transcripts from MT-1ψb is that this pseudogene directs the synthesis of an mRNA which is unstable since it lacks the proper consensus sequence (TGTCGTTT at position 1034, Fig. 2). To test this possibility we constructed a fusion gene between the 5′ regulatory region of MT-1ψb and the CAT gene. We called this construct pMT(ψb)cat; it contains 1.1 kb of sequence from the 5′ end of the MT-1ψb pseudogene, followed by the CAT gene. This construct is analogous to pMT(p+)cat, in which 126 bp of MT regulatory sequence (in common with the pseudogene sequence) are fused to the CAT gene coding sequence.

We transfected 293 cells with pMT(ψb)cat, as well as with pSV2cat and $pA_{10}cat_2$, and assayed cell extracts for CAT enzyme activity 48 h later. Cells transfected with pMT(ψb)cat showed surprisingly high levels of CAT activity in the absence of inducer and a twofold increase in CAT activity after a 6-h exposure to 5 μM $CdCl_2$ (Fig. 10B). We conclude that this pseudogene regulatory region is both transcriptionally active and responsive to metal ion regulation. An examination of MT-ψb sequences shows several "TATA" sequences near the 5′ *Hin*dIII end (sequences not shown; available on GenBank) which in 293 cells may be acting as cryptic promoter sites and producing a background level of transcription. After transfection of HeLa cells with this construct we found uninduced cells exhibited CAT activity comparable to $pA_{10}cat_2$; $CdCl_2$-treated cells showed a fivefold stimulation of CAT enzyme activity (data not shown). This difference between HeLa and 293 cells may represent the lack of activation in HeLa cells of cryptic promoter sites that we suspect contribute to transcriptional activity in the 293 cell transfection experiments.

DISCUSSION

Both the human (38) and rat metallothionein genes belong to multigene families. In both cases, many of these sequences are due to the presence of processed pseudogenes. In Fig. 11 we summarize the relationship between the MT-1 structural gene and pseudogenes MT-1ψa, MT-1ψb, and MT-1ψc. Most processed pseudogenes show a short (8 to 14 bp) direct repeat flanking the insertion site (see Vanin, in press, for review). Insertion of cDNA copies of mRNAs, like that of transposable elements, may proceed into a site in which the recipient DNA has been cut in a staggered fashion. Insertion then duplicates this site to produce a short direct repeat. All three MT-1 pseudogenes may have arisen by this mechanism. In the case of MT-1ψa, the duplication of the polyadenylation terminator at the 5′ end of this inserted sequence suggests that the site of insertion contained an A-rich region. The formation of pseudogene MT-1ψc (which has lost its 3′ sequences) by reverse transcription of an MT-1 mRNA can be deduced from the observation that there is complete homology between the 5′ end of this pseudogene and the 5′ end of the MT-1 transcription unit. This homology does not extend 5′ to the MT-1 transcription start site (Fig. 2).

Metallothioneins are ubiquitous in the tissues of the body (17). The expression of these genes in germline tissue would permit processed pseudogenes to form by the reverse transcription of message (Vanin, in press). Numerous examples of such pseudogenes have now been documented. Indeed, in one case, that of arginino-succinate synthetase, 14 pseudogenes were found in addition to the normal structural gene (21). It is of interest that the mouse, in contrast to the rat, does not contain any detectable MT-1 pseudogenes (6, 31). One explanation for this difference may be a possible species difference in the expression of MT genes in germline cells.

The reverse transcription of mRNA to produce a processed pseudogene would not normally be expected to create a new functional gene, because the regulatory elements 5′ to the normal transcription initiation point would be lost. The existence of upstream transcription initiation sites, recognized by either polymerase II or III (23), could allow the formation of mRNA species which include the promoter sequences of structural genes, and which could be processed, polyadenylated, and reverse transcribed to give rise to new genes which contain transcription units with their 5′ regulatory elements preserved (see review on pseudogenes in Vanin et al.). Pseudogene MT-1ψb is an example of such a new gene. However, we cannot detect MT-1ψb transcripts, either from Northern analysis or from S1 protection analysis of RNA from transfection experiments. One explanation for this is the lack of 3′ sequences required for efficient formation of $polyA^+$ mRNA 3′ termini. Another explanation for our inability to detect transcripts of MT-1ψb is a reduction in the frequency of transcription due to a loss of 5′ promoter sequences. We compared the construct pMT-1(p+)cat containing 126 bp of sequence 5′ of transcription initiation (Fig. 10A) with an analogous construct containing 920 bp of the MT gene 5′ of transcription initiation fused to CAT. We found a 15-fold increase in both the basal level and the cadmium-induced level of CAT activity in this latter construct when transfected into 2MO cells (rat hepatoma cell line; unpublished observations). The shortened promoter region in pseudogene MT-1ψb (142 bp of 5′ homologous sequence) may result in an attenuation in promoter activity and a reduced level of accumulated transcripts.

The observation that the "TATA" sequence 5′ to the MT-1 promoter both transcribes and may also be regulated by metal ion was unexpected. There are no sequences within this promoter region which resemble the consensus sequences found in the MT-1 promoter. We are testing the hypothesis that these metal regulatory consensus sequences may exert an effect on adjacent promoters, similar to observations made with enhancer sequences (see reference 29 for a review), or with the glucocorticoid regulatory element (13), the mouse MT-1 metal regulatory element (11, 12) or the human MT-2a metal regulatory element (35) adjacent to the thymidine kinase gene. Studies with fusion genes, utilizing MT regulatory sequences, as well as with defective genes such as MT-1ψb, will allow us to understand better the mechanisms underlying the regulation of gene transcription and to explore the specific manner in which metal ions or other agents regulate transcription of metallothionein genes.

ADDENDUM

Soares et al. (63) have recently demonstrated that preproinsulin I in rats and mice is a retropseudogene resulting from a truncated transcript originating from an upstream

TATA box (−804 bp). This upstream promoter is functional in a transcription system in vitro, but does not produce detectable transcripts in vivo. The 5′ and 3′ ends of this gene form 41-bp direct repeats which eliminate the consensus sequence 3′ of the polyadenylation site (45). Unlike pseudogene MT-1ψb however, rat and mouse preproinsulin I and preproinsulin II are almost equally expressed in vivo.

ACKNOWLEDGMENTS

We thank Tom Ganz and George Neumann for help with Sanger sequencing; Garry Miyada and Bruce Kohorn for help in the use of riboprobe vectors, and Mark Mercola for helping us with the CAT assay. We acknowledge the expert technical assistance of Frances Holzer, Sandy Wong, and Greg Bristol, and thank Sharon Morse for typing this manuscript.

This work was supported by contract DE AC03 76 SF00012 from the Department of Energy; Public Health Service (H.R.H.), grant AM31183A from the National Institutes of Health (R.D.A.), and BIONET grant 1 U41 RR-01685-02. S.J.T. is a predoctoral fellow supported by training grant GM 07185 from the National Institutes of Health.

LITERATURE CITED

1. **Albertini, A. M., M. Hofer, M. P. Calos, and J. H. Miller.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell **29**:319–328.
2. **Andersen, R. D., B. W. Birren, T. Ganz, J. E. Piletz, and H. R. Herschman.** 1983. Molecular cloning of the rat metallothionein-1 (MT-1) mRNA sequence. DNA **2**:15–22.
3. **Andersen, R. D., J. E. Piletz, B. W. Birren, and H. R. Herschman.** 1983. Levels of metallothionein messenger RNA in foetal, neonatal and maternal rat liver. Eur. J. Biochem. **131**:497–500.
4. **Andersen, R. D., W. P. Winter, J. J. Maher, and I. A. Bernstein** 1978. Turnover of metallothioneins in rat liver. Biochem. J. **174**:327–338.
5. **Andrews, G. K., E. D. Adamson, and L. Gedamu.** 1984. The ontogeny of expression of murine metallothionein: comparison with the α-fetoprotein gene. Develop. Biol. **103**:294–303.
6. **Beach, L. R., and R. D. Palmiter.** 1981. Amplification of the metallothionein-I gene in cadmium-resistant mouse cells. Proc. Natl. Acad. Sci. USA **78**:2110–2114.
7. **Benoist, C., and P. Chambon.** 1981. In vivo sequence requirements of the SV40 early promoter region. Nature (London) **290**:304–315.
8. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
9. **Breathnach, R., C. Benoist, K. O'Hare, F. Gannon, and P. Chambon.** 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc. Natl. Acad. Sci. USA **75**:4853–4857.
10. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
11. **Brinster, R. L., H. Y. Chen, R. Warren, A. Sarthy, and R. D. Palmiter.** 1982. Regulation of metallothionein-thymidine kinase fusion plasmids injected into mouse eggs. Nature (London) **296**:39–42.
12. **Carter, A. D., B. K. Felber, M. Walling, M.-F. Jubier, C. J. Schmidt, and D. H. Hamer.** 1984. Duplicated heavy metal control sequences of the mouse metallothionein-I gene. Proc. Natl. Acad. Sci. USA **81**:7392–7396.
13. **Chandler, V. L., B. A. Maler, and K. R. Yamamoto.** 1983. DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. Cell **33**:489–499.
14. **Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon.** 1980. Promoter sequences of eukaryotic protein-coding genes. Science **209**:1406–1414.
15. **Dayhoff, M. O.** 1969. Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Spring, Md.
16. **Durnam, D. M., J. S. Hoffman, C. J. Quaife, E. P. Benditt, H. Y. Chen, R. L. Brinster, and R. D. Palmiter.** 1984. Induction of mouse metallothionein-I mRNA by bacterial endotoxin is independent of metals and glucocorticoid hormones. Proc. Natl. Acad. Sci. USA **81**:1053–1056.
17. **Durnam, D. M., and R. D. Palmiter.** 1981. Transcriptional regulation of the mouse metallothionein-I gene by heavy metals. J. Biol. Chem. **256**:2268–2272.
18. **Durnam, D. M., and R. D. Palmiter.** 1984. Induction of metallothionein-I mRNA in cultured cells by heavy metals and ioidoacetate: evidence for gratuitous inducers. Mol. Cell. Biol. **4**:484–491.
19. **Etzel, K. R., S. G. Shapiro, and R. J. Cousins.** 1978. Regulation of liver metallothionein and plasma zinc by the glucocorticoid dexamethasone. Biochem. Biophys. Res. Commun. **89**:1120–1126.
20. **Failla, M. L., and R. J. Cousins.** 1978. Zinc accumulation and metabolism in primary cultures of adult rat liver cells: regulation by glucocorticoids. Biochim. Biophys. Acta **543**:293–304.
21. **Freytag, S. O., H.-G. O. Bock, A. L. Beaudet, and W. E. O'Brien.** 1984. Molecular structures of human argininosuccinate synthetase pseudogenes. Evolutionary and mechanistic implications. J. Biol. Chem. **259**:3160–3166.
22. **Friedman, R. L., and G. R. Stark.** 1985. α-Interferon-induced transcription of HLA and metallothionein genes containing homologous upstream sequences. Nature (London) **314**: 637–639.
23. **Galli, G., H. Hofstetter, and M. L. Birnstiel.** 1981. Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature (London) **294**:626–631.
24. **Garoff, H., and W. Ansorge.** 1981. Improvements of DNA sequencing gels. Anal. Biochem. **115**:450–457.
25. **Gaynor, R. B., D. Hillman, and A. Berk.** 1984. Adenovirus early region 1A activates transcription of a nonviral gene introduced into mammalian cells by infection or transfection. Proc. Natl. Acad. Sci. USA **81**:1193–1197.
26. **Glanville, N., D. M. Durnam, and R. D. Palmiter.** 1981. Structure of mouse metallothionein-I gene and its mRNA. Nature (London) **292**:267–269.
27. **Graham, F. L., J. Smiley, W. C. Russell, and R. J. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. Gen. Virol. **36**:59–72.
28. **Gross-Bellard, M., P. Oudet, and P. Chambon.** 1973. Isolation of high molecular weight DNA from mammalian cells. Eur. J. Biochem. **36**:32–36.
29. **Gruss, P.** 1984. Magic enhancers. DNA **3**:1–5.
30. **Hager, L. J., and R. D. Palmiter.** 1981. Transcriptional regulation of mouse liver metallothionein-I gene by glucocorticoids. Nature (London) **291**:340–342.
31. **Herschman, H. R.** 1985. A 12-*O*-tetradecanoyl phorbol-13 acetate (TPA) nonproliferative variant of 3T3 cells is resistant to TPA-enhanced gene amplification. Mol. Cell. Biol. **5**:1130–1135.
32. **Kägi, J. H. R., and M. Nordberg, ed.** 1979. Metallothionein, proceedings of the first international meeting on metallothionein and other low molecular weight metal-binding proteins. Zürich, 17–22 July 1978. Birkhauser Verlag, Boston, Mass.
33. **Kanehisa, M. I.** 1982. Los Alamos sequence analysis package for nucleic acids and proteins. Nucleic Acids Res. **10**:183–196.
34. **Karin, M., A. Haslinger, H. Holtgreve, G. Cathala, E. Slater, and J. D. Baxter.** 1984. Activation of a heterologous promoter in response to dexamethasone and cadmium by metallothionein gene 5′-flanking DNA. Cell **36**:371–379.
35. **Karin, M., A. Haslinger, H. Holtgreve, R. I. Richards, P. Krauter, H. M. Westphal, and M. Beato.** 1984. Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-II_A gene. Nature (London) **308**:513–519.
36. **Karin, M., and H. R. Herschman.** 1979. Dexamethasone stimulation of metallothionein synthesis in HeLa cell cultures. Science **204**:176–177.
37. **Karin, M., and H. R. Herschman.** 1980. Glucocorticoid hormone receptor mediated induction of metallothionein synthesis

in HeLa cells. J. Cell. Physiol. **103**:35–40.
38. **Karin, M., and R. I. Richards.** 1982. Human metallothionein genes; primary structure of the metallothionein-II gene and a related processed gene. Nature (London) **299**:797–802.
39. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
40. **Lewin, B.** 1980. Alternatives for splicing: recognizing the ends of introns. Cell **22**:324–326.
41. **Li, W-H., T. Gojobori, and M. Nei.** 1981. Pseudogenes as a paradigm of neutral evolution. Nature (London) **292**:237–239.
42. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
43. **Mayo, K. E., R. Warren, and R. D. Palmiter.** 1982. The mouse metallothionein-I gene is transcriptionally regulated by cadmium following transfection into human or mouse cells. Cell **29**:99–108.
44. **McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith.** 1984. The distal transcription signals of the herpes virus tk gene share a common hexonucleotide control sequence. Cell **37**:253–262.
45. **McLauchlan, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**:1347–1368.
46. **Mercola, M., J. Goverman, C. Mirell, and K. Calame.** 1985. Immunoglobulin heavy chain enhancer requires one or more tissue-specific factors. Science **227**:266–270.
47. **Miyata, T., and T. Yasunaga.** 1981. Rapidly evolving mouse α-globin-related pseudo gene and its evolutionary history. Proc. Natl. Acad. Sci. USA **78**:450–453.
48. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
49. **Nordberg, G. F.** 1972. Cadmium metabolism and toxicity: experimental studies on mice with special reference to the use of biological materials as indices of retention and the possible role of metallothionein in transport and detoxification of cadmium. Environ. Physiol. Biochem. **2**:7–37.
50. **Nordheim, A., and A. Rich.** 1983. The sequence $(dC\text{-}dA)_n \times (dG\text{-}dT)_n$ forms left-handed Z-DNA in negatively supercoiled plasmids. Proc. Natl. Acad. Sci. USA **80**:1821–1825.
51. **Oh, S. H., J. T. Deagan, P. D. Whanger, and P. H. Weswig.** 1978. Biological function of metallothionein. Part 5. Its induction in rats by various stresses. Am. J. Physiol. **234**:E282–E285.
52. **Piletz, J. E., R. D. Andersen, B. W. Birren, and H. R. Herschman.** 1983. Metallothionein synthesis in foetal, neonatal and maternal rat liver. Eur. J. Biochem. **131**:489–495.
53. **Pitot, H., C. Peraino, P. A. Morse, and V. R. Potter.** 1964. Hepatomas in tissue culture compared with adapting liver *in vivo*. Natl. Cancer Inst. Monogr. **13**:229–246.
54. **Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey.** 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. USA **79**:4298–4302.
55. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′-Non-coding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
56. **Rice, N. R.** 1972. Change in repeated DNA in evolution, p. 44–79. *In* H. H. Smith (ed.), Evolution of genetic systems. Gordon and Breach, New York.
57. **Richards, R. I., A. Heguy, and M. Karin.** 1984. Structural and functional analysis of the human metallothionein-I_A gene: differential induction by metal ions and glucocorticoids. Cell **37**:263–272.
58. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
59. **Ryden, L., and H. R. Deutsch.** 1978. Preparation and properties of the major copper binding component in human fetal liver: its identification as metallothionein. J. Biol. Chem. **253**:519–524.
60. **Sanger, F., A. R. Coulsen, B. G. Barrell, A. J. H. Smith, and B. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA-sequencing. J. Mol. Biol. **143**:161–178.
61. **Searle, P. F., B. L. Davison, G. W. Stuart, T. M. Wilkie, G. Norstedt, and R. D. Palmiter.** 1984. Regulation, linkage, and sequence of mouse metallothionein-I and II genes. Mol. Cell. Biol. **4**:1221–1230.
62. **Shen, S-H., J. L. Slightom, and O. Smithies.** 1981. A history of the human fetal globin gene duplication. Cell **26**:191–203.
63. **Soares, M. B., E. Schon, A. Henderson, S. K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin, and A. Efstratiadis.** 1985. RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Mol. Cell. Biol. **5**:2090–2103.
64. **Sobocinski, P. A., W. J. Canterbury, C. A. Mapes, and R. E. Dinterman.** 1978. Involvement of hepatic metallothioneins in hypozincemia associated with bacterial infection. Am. J. Physiol. **234**:E399–E406.
65. **Southern, E.** 1975. Detection of specific sequences among fragments separated by gel electrophoresis. J. Mol. Biol. **998**:503–517.
66. **Staden, R.** 1980. A new computer program for the storage and manipulation of DNA gel reading data. Nucleic Acids Res. **8**:3673–3694.
67. **Stuart, G. W., P. F. Searle, H. Y. Chen, R. L. Brinster, and R. D. Palmiter.** 1984. A 12-base-pair DNA motif that is repeated several times in metallothionein gene promoters confers metal regulation to a heterologous gene. Proc. Natl. Acad. Sci. USA **81**:7318–7322.
68. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
69. **Van der Eb, A. S., and F. L. Graham.** 1980. Assay of transforming activity of tumor virus DNA. Methods Enzymol. **65**:826–839.
70. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
71. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.