The Human and Mouse GATA-6 Genes Utilize Two Promoters and Two Initiation Codons*

(Received for publication, June 11, 1999, and in revised form, August 30, 1999)

Alison Brewer‡, Christopher Gove§, Andy Davies¶, Claire McNulty§, Dalna Barrow‡,
Manoussos Koutsourakis‡, Farzin Farzaneh†, John Pizsey§, Adrian Bomford¶,
and Roger Patient§**

From the ‡Department of Molecular Medicine, The Rayne Institute, GKT, 123 Coldharbour Lane, London SE5 9NU, United Kingdom, the §Developmental Biology Research Centre, The Randall Institute, King’s College London, 26-29 Drury Lane, London WC2B 5RL, United Kingdom, ¶Institute of Liver Studies, GKT, Bessemer Rd., London SE5 9PJ, United Kingdom, and §Erasmus University, Medical Genetics Centre, Department of Cell Biology and Genetics, 3000 DR Rotterdam, The Netherlands

GATA-6 has been implicated in the regulation of myocardial differentiation during cardiogenesis. To determine how its expression is controlled, we have characterized the human and mouse genes. We have mapped their transcriptional start sites and demonstrate that two alternative promoters and 5’ noncoding exons are utilized. Both transcript isoforms are expressed in the same tissue-specific and developmental stage-specific pattern, and their ratio appears similar wherever examined. The more upstream noncoding exon showed a substantial degree of homology between the two mammalian species, suggesting a conserved regulatory function. Moreover, in transcription assays we show that elements within this exon act to promote its transcription. Positive regulatory elements that effect transcription from the more downstream exon were not apparent in this assay, revealing a regulatory distinction between the two promoters. We also demonstrate alternative initiator codon usage in both the human and mouse GATA-6 genes. Both isoforms of the protein are synthesized in vitro regardless of which 5’ noncoding exon is present in the RNA, although the larger protein has greater transcriptional activation potential in transfection assays. Thus, GATA-6 function in the cell is controlled by a complex interplay of transcriptional and translational regulation.

Members of the GATA family of zinc finger transcription factors play essential roles in the regulation of cellular differentiation during vertebrate development. Six distinct members have so far been identified, which can broadly be classified into two subfamilies. GATA-1, -2, and -3 are all found in specific hematopoietic cells and a number of ectodermal derivatives (1–3). By contrast, GATA-4, -5, and -6 show partially overlapping expression patterns in cardiac and endodermally derived cells (4). GATA-6 is one of the earliest and most crucial GATA genes to be expressed during embryogenesis. In mouse embryos, zygotic GATA-6 expression can be detected from as early as the blastocyst stage at 3.5 days postcoitum (dpc) within a proportion of cells of the inner cell mass. Moreover, targeted mutagenesis of the mouse GATA-6 gene is lethal to developing embryos at 5.5 dpc (shortly after implantation and before gastrulation), due to gross retardation and abnormal primitive endoderm differentiation (5). This is the earliest lethality reported for any single GATA factor. In addition, early visceral endodermal markers such as HNF4 and GATA-4 are not expressed in GATA-6-deficient mice, indicating that GATA-6 lies upstream of these other transcription factors in a transcriptional cascade that regulates differentiation of the visceral endoderm (6). After gastrulation, mouse GATA-6 is expressed in the parietal endoderm and subsequently in the mesoderm and endoderm that contribute to the heart and gut, respectively (5).

In Xenopus, we have also detected expression of GATA-6 in the endoderm and mesoderm of early gastrulae, including the presumptive heart mesoderm (7). mRNA levels in precardiac cells normally decrease prior to terminal differentiation; however, by artificially maintaining high levels of GATA-6 transcripts, we could delay the progression of the myocardial differentiation program. Moreover, as these elevated levels of GATA-6 within the embryos were allowed to decay, differentiation could then progress and resulted in an enlarged myocardium. Thus, GATA-6 may act to maintain early cardiac progenitors in a precursor state, and its absolute cellular levels (possibly in relation to other GATA factors) may provide an important regulatory cue in the differentiation program. It is interesting to note that overexpression of GATA-2 in erythroid progenitors has been similarly shown to inhibit terminal differentiation (8). Thus, GATA-2 and GATA-6 may play similar roles in erythroid and cardiogenic cells, respectively.

To begin to elucidate the molecular mechanisms that regulate the temporal and spatial expression of GATA-6, we have isolated and characterized the mouse and human GATA-6 genes. We find that by contrast to chicken GATA-6, which is transcribed from a single promoter (9), the mammalian genes both contain two alternative noncoding upstream exons, transcribed from two distinct promoters, as reported previously for

---

* This work was supported by the British Heart Foundation and by studentships from King’s College London (to C. M.) and from Joint Research Committee, King’s Healthcare/King’s College School of Medicine and Dentistry (to A. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) AJ243146 and AJ243149.

** To whom correspondence should be addressed. Present address: Institute of Genetics, University of Nottingham, Queen’s Medical Centre, Nottingham, NG7 2UH. E-mail: roger.patient@nottingham.ac.uk.
GATA-6 Gene Organization, Expression, and Translation

some other members of the GATA family (2, 10, 11, 12). By a combination of RNase protection analysis, exon-specific in situ hybridization, and RT-PCR, we have not detected any difference in the spatial expression patterns of the two GATA-6 transcript isoforms. In addition, in all of the tissues and developmental stages we have examined, significant changes in the ratio of the two transcripts have not been seen. We note the exceptionally high degree of sequence conservation of the more upstream, 5’ noncoding exon, between the GATA-6 genes of the two mammalian species. This is suggestive of an evolutionarily conserved regulatory function for this (transcribed) region, which is supported by transient transfection assays. This region stimulates transcription only from the upstream exon, suggesting that, although the two 5’ exons are frequently co-expressed, their corresponding two promoters are independently regulated. Finally, we show that the mammalian GATA-6 genes encode a longer polypeptide than has previously been reported (13–15), which initiates from an upstream in-frame methionine codon. We demonstrate that in vitro both the long and the short forms of GATA-6 can be produced, with the longer protein showing significantly more transactivation activity in transfection assays. Thus, the mammalian GATA-6 genes exhibit a complexity in their regulation, at both the levels of transcription and translation.

EXPERIMENTAL PROCEDURES

Characterization of Human and Mouse GATA-6 Genomic Clones—Genomic GATA-6-containing clones (kindly provided by F. Grosveld) were mapped by standard protocols (16). Restriction fragments that hybridized to human (14) or mouse (kind gift from F. Grosveld) cDNA clones were cloned into pBluescript II SK+ (Stratagene) and were sequenced using a combination of gene-specific and universal M13 sequencing primers. The positions of exons were determined by the polymerase chain reaction (PCR), and delineation of intron/exon boundaries was accomplished by DNA sequencing. Sequence analysis was performed using the University of Wisconsin GCG suite of program for molecular biological analysis (17).

5’-RACE and RT-PCR—mRNA was isolated using a SV Total RNA Isolation kit (Promega). 5’-RACE reactions were performed using a CAP Fider kit (CLONTECH) and the following primers: for human Ia, an antisense primer directed against nucleotides 6–10 of the published human GATA-6 CDNA sequence; for mouse exon Ia, the oligonucleotide representing the complement to nucleotides 172–193 (Fig. 1C) was used; and for mouse Ib, the antisense primer GC-GCCGAAACGCTTCGGCAG, complementary to sequences within the 5′ region of exon II. The resultant PCR fragments were either sequenced directly using the primers described or were cloned into the pCR 2.1 vector with a TA cloning kit (Invitrogen) and sequenced using universal primers.

For RT-PCRs, 1 μg of total RNA from various tissues as described was reverse transcribed into cDNA with Maloney murine leukemia virus reverse transcriptase (Stratagene) according to the manufacturer’s specifications, using random 10-mer. These cDNAs were then used as templates in PCRs. Sense primers for human and mouse exons Ia and Ib, in addition to an antisense primer for human exon Ia, are as follows: human and mouse exon Ia sense primer, GCTGTTGTTTAGGGCTGC; human Ib sense primer, GCGGTTGCTGTGTGGAG-AC; mouse exon Ib sense primer, GAGAGTGGTGAGAGTGGAG; human exon II antisense primer, AAGGATGGACGAGGTTAG; mouse exon II antisense primer used is as described above. After denaturation at 95 °C for 4 min, 40 PCR cycles were typically performed and consisted of 94 °C for 1 min, 60 °C for 30 s, and 72 °C for 1 min, using SuperTaq thermostable DNA polymerase (HT Biotechnology Ltd.) in the manufacturer’s supplied buffer. As above, the resultant PCR fragments were either sequenced directly using the primers described or were cloned into the pCR 2.1 vector with a TA cloning kit (Invitrogen) and sequenced using universal primers.

RNase Protection Analysis—The RNase protection probe used to map the mouse exon Ia CAP site was a 292-bp BamHI–XhoI genomic fragment cloned into the polylinker of Bluescript II SK+ (Stratagene). After linearizing with BamHI, antisense RNA was transcribed with T7 polymerase (Promega). The mouse exon Ib CAP site probe was a 480-bp genomic PCR fragment, with XhoI and BamHI linkers at its 5′ and 3′ ends respectively, which spanned the exon Ib CAP site (Fig. 2). This was cloned into pBluescript II SK+ (Stratagene), linearized with XbaI, and transcribed with T7 polymerase (Promega).

The RNase protection probes that spanned the exon Ia-II or exon Ib-II boundaries were constructed by overlapping PCR. A CDNA fragment representing nucleotides 31–221 of exon II (Fig. 1) with an artificial polylinker was generated at its 3′ end and used to transcribe a mouse GATA-6 cDNA clone as template. This was extended in an overlapping PCR to cDNA sequences representing either position 172 of exon Ia to position 84 of exon II or position 250 of exon Ib to position 84 of exon II (Fig. 1), which were similarly generated with an artificial XbaI site at their 5′ ends by PCR using 5′ cDNA clones as templates. The resulting products were digested with XbaI and PstI and cloned into the appropriate sites within pBluescript II SK+ to generate Mia-II and Mil-II, respectively. Plasmids were linearized with XbaI and transcribed with T7 polymerase. For the glycolaldehyde-3-phosphate dehydrogenase control, the probe described previously (18) was digested with AccI, and the resultant 132-bp fragment was subcloned into pGem3Z. This was subsequently digested with HinIII and transcribed with T7 polymerase (Promega).

Protection assays were performed as described previously (7). After autoradiography, the intensities of protected species and their respective glycolaldehyde-3-phosphate dehydrogenase signals were quantitated using a GS 250 Molecular Imager (Bio-Rad).

In Situ Hybridization—The mouse pan-GATA-6 probe was a 1.5-kb EcoRI fragment of human (14) and mouse (kind gift from F. Grosveld) cDNA fragment representing nucleotides 31–221 of exon II (Fig. 1), which were similarly generated with an artificial XbaI site at their 5′ ends by PCR using 5′ cDNA clones as templates. The resulting products were digested with XbaI and PstI and cloned into the appropriate sites within pBluescript II SK+ (Stratagene), linearized with XbaI, and transcribed with T7 polymerase. Probes for mouse Ia and Ib exon-specific in situ hybridization were PCR fragments, linked with XbaI and BamHI at their 5′ and 3′ ends, respectively, that represented either the whole of exon Ia or exon Ib, which were generated using a 1.4-kb BamHI genomic subclone spanning both exons as template (see Fig. 1A). The fragments were cloned into the polylinker of pBluescript II SK+ (Stratagene), linearized with XbaI, and transcribed with T7 polymerase (Promega). The transcription mix included Digoxygenin-11-UTP (Roche Molecular Biochemicals). This comprised 1/3 of the total UTP for the pan-GATA-6 probe and 1/2 of the total UTP for the exon-specific probes. For in situ hybridization on sections, mouse embryos were fixed overnight in 4% paraformaldehyde, dehydrated in washes of 25, 50, 75, and 100% methanol (5 min each); and washed in ethanol (2 × 30 min), xylene (1 × 45 min), and finally paraﬃn wax (3 × 20 min, 60 °C). During the final paraﬃn wash they were orientated and then left overnight to set. 20-μm sections were mounted on treated slides (19), and in situ hybridizations were performed according to a protocol adapted from the whole mount protocol previously described (20). Prehybridization was performed in an aqueous solution (2 × 10 min), washed in ethanol (2 × 5 min), and treated with 6% hydrogen peroxide in methanol (3 h). Prehybridization and hybridization steps were as described (20) as were posthybridization washes, except for the omission of the RNase step. The slides were blocked with 2% Roche Molecular Biochemicals blocking reagent in maleic acid buffer (2 h) and the immunodetection using an alkaline phosphatase conjugated anti-dig antibody (Roche Molecular Biochemicals) was performed as described (20). The color was developed using 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium as the substrate in a polyvinyl acetate-containing buffer (21).

Construction of Mouse GATA-6-CAT Reporter Plasmids—Genomic restriction fragments of 0.94 kb (KpnI–NhelI), 1.15 kb (KpnI–XhoI), or 0.2 kb (Nhel–XhoI) that span the mouse GATA-6 exon Ia were cloned into the appropriate sites within the polylinker of the pCAT-Basic promotorless vector (Promega) to yield MIa-CAT, MIIa + NX-CAT, and NX-CAT, respectively. To construct MXb-CAT and MXb + NX-CAT, a PCR fragment was generated with a BamHI-linked 3′ end. The template for the PCR product was a 1.4-kb BamHI genomic fragment that spans exons Ia and Ib (see Fig. 1A), subcloned into pBluescript II SK+ (Stratagene). Using the M13 universal forward primer and an oligonucleotide of sequence GAGGTTGATCTGTGGACGGTGCACATGGAG, which contains a BamHI site, adjacent to sequences which are complementary to a region within exon Ib (mouse nucleotides 1109–1131 in Fig. 1C), a fragment of approximately 1.2 kb was generated. This was digested with BamHI and cloned into the BglII site of pCAT-Basic. Subsequent digestion and religation of this construct with either XhoI or PstI yielded MXb-CAT and MXb + NX-CAT, respectively.

WRL Cell Culture and Transient Transfection CAT Assays—The human fetal hepatocyte cell line, WRL, was grown in Dulbecco’s minimal essential medium and 10% fetal calf serum. Transient transfections were performed in duplicate using TransFast reagent (Promega). Cells were transfected in 25-cm2 flasks, at 50% confluence, according to the manufacturer’s instructions, with 5 μg of test CAT-reporter construct and 200 ng of control pBS-p-gal (Promega). Transfected cells...
were harvested after 48 h and were lysed with 1 ml of reporter lysis buffer. CAT assays were performed on aliquots (78 µl) of the lysates by scintillation counting. β-Galactosidase activity in 0.15-ml aliquots of the same lysates was determined using an enzyme assay system kit (Promega), exactly as described (16). HepG2 Cell Culture and Transient Transfection Luciferase Assays—The human HepG2 hepatoma cell line was transfected using the calcium phosphate-mediated cell precipitation method (22). 0.4 × 10^6 cells were plated in 60-mm six-well tissue culture plates and grown overnight in 2 ml of Eagle's minimal essential medium with 10% fetal calf serum. Transfections were performed with 0.1 µg of expression plasmid (see below for details of constructs) or 1 µg of GATA-Luc reporter construct (23) or mutGATA-Luc negative reporter construct (both kind gifts of M. Yamamoto), and 0.5 µg of pCMVβ (CLONTECH) with Protéroga's luciferase assay system. 50-µl aliquots of the extract was assayed for luciferase activity in a Lucyl luminometer (Rosys Anthos). Transfected cells were harvested and lysed with 250-µl of the extract was added. Transfected cells were harvested and lysed with 250-µl passive lysis buffer (Promega) after 48 h. 20 µl of the extract was assayed for luciferase activity in a Luxel luminometer (Rosys Anthos) using Promega's luciferase assay system. 50-µl aliquots of the extract were assayed for β-galactosidase activity in 96-well plates using the Promega β-galactosidase enzyme assay system; yellow reaction products were quantified using an MRX plate reader (Dynex). The luciferase activities were normalized for transfection efficiency against the β-galactosidase activity.

In Vitro Transcription and Translation—Human constructs were all derived by PCR using a linearized, full-length human GATA-6 cDNA (14), cloned into pUC21 (pMALT-UTR) as template. PCR was performed with AmpliTag Gold (Perkin-Elmer) using the following cycle conditions: 94 °C for 10 min and then 35 cycles of 94 °C for 30 s and 68 °C for 30 s with a final extension reaction of 72 °C for 10 min. 1% Me2SO was added to the reactions to reduce the adverse effects of secondary structure in the template DNA. The primers used are defined in Table I. All numbers refer to our published cDNA sequence (14). The MALT-ORF and MYQ-ORF expression constructs, comprising 5′ deletions of the cDNA clone (pMALT-UTR) were constructed by subcloning fragments produced using either the MALT-BAM or MYQ-BAM 5′ primers (which included 5′ SacI restriction sites) together with the downstream primer 13. The PCR products were restricted with BamHI and BstXI and ligated into the vector-containing fragment of BamHI/SacI-restricted pMALT-UTR. The resultant plasmids were restricted with BamHI and XhoI, and the inserts were subcloned into the appropriate sites within the multipurpose expression vector pCS-2+ vector (24, 25) to generate MALT-ORF and MYQ-ORF. The MYQ-20 deletion construct was made using the MYQ-BAM 5′ primer and the MYQ-20 3′ primer, which incorporates a SmaI site. The PCR product was restricted with BamHI and SmaI and ligated into BamHI/SmaI-restricted MALT-ORF. MYQ-77 was also made in this way using the MYQ-BAM 5′ primer and the MYQ-77 3′ primer. The MALT deletion constructs (MALT-18 and MALT-146) were made using the 5′ MALT-BAM primer and the MALT-18 or MALT-146 3′ primers that include either a HincII or a SmaI restriction site, respectively. The PCR products were restricted with the appropriate enzymes and ligated into BamHI/SmaI-restricted MALT-ORF.

Site-directed mutagenesis was performed using Promega's Altered Sites II system with the insert from the MALT-ORF expression construct subcloned into the BamHI and KpnI sites of the pALTER-1 vector. The single-stranded pALTER-1-MALT-ORF DNA was prepared from phage for subsequent mutation reactions. Mutation of the internal ATG to TTG was performed by carrying out the PCR with the M147L oligonucleotide (Table I) that anneals to the antisense strand.

Mouse clones were constructed by overlapping PCR. cDNA fragments comprising nucleotides 31 of exon II to position 8 of exon III (Fig. 1) were generated adjacent to a TGA stop codon and artificial KpnI site by PCR using a mouse GATA-6 cDNA as template. This was extended by overlapping PCR to fragments representing either position 37 of exon Ia to position 84 of exon II or position 1 of exon Ib to position 84 of exon II, also generated by PCR with an artificial XbaI site at their 5′ ends using mouse GATA-6 cDNA clones as templates. The resulting extended products were digested with XbaI and KpnI and cloned into the appropriate sites within pBluescript II KS+ to generate MxIA-I and MXb-II, respectively.

Reactions were performed with the SP6 or T7 polymerase TNT kits (Promega) for human and mouse constructs, respectively, using purified plasmid DNA (Qiagen, Chatsworth CA), and [35S]methionine (Amersham Pharmacia Biotech). Translation products were separated on denaturing SDS-polyacrylamide gels and visualized by autoradiography. Western Blotting—Protein samples were electrophoresed on 10% polyacrylamide gels (26) and electrophoibrated onto nitrocellulose membrane (Amersham Pharmacia Biotech ECL membrane) essentially as described (27). The membranes were exposed directly to x-ray film for the detection of incorporated [35S]methionine. The membranes were then blocked with 2% milk powder in PBS, 0.1% Tween, and GATA-6 protein was detected with a 1:2000 dilution of goat anti-human GATA-6 antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) and a 1:5000 dilution of rabbit anti-goat horseradish peroxidase-conjugated secondary antibody (Dako). Horseradish peroxidase activity was detected by ECL (Amersham Pharmacia Biotech) using the standard protocol. A clear sheet of film was placed between the nitrocellulose membrane and the photographic film to absorb 35S radiation during ECL detection.

RESULTS

The Genomic Structure of the Human and Mouse GATA-6 Genes

Cosmid clones were isolated from both mouse and human genomic libraries using mouse or human GATA-6 cDNA fragments as hybridization probes (kind gifts of F. Grosveld) and were characterized by a combination of restriction mapping and DNA sequencing. Partial intron-exon boundary maps derived from these clones are shown in Fig. 1A. Exons Ia and Ib are untranslated (see below), and the protein coding regions of the genes are contained within at least six exons (exons II–VII), the two zinc fingers being encoded separately on exons III and IV. The existence of exon VII is deduced from the presence of both additional coding and 3′-untranslated sequences in the cDNAs that are not present in exon VI. The consensus splice donor and acceptor sites surrounding exons II and III in the mouse gene and IV, V, and VI in humans are shown in Fig. 1B.

The Human and Mouse GATA-6 Genes Contain Two Alternative Noncoding 5′ Exons

Identification of Exon Ia—We have previously reported the cloning of a human GATA-6 cDNA from a fetal heart library (14). By Southern blot and sequence comparison with human genomic clones from the GATA-6 locus, we found the gene to
contain a noncoding 5' exon, in common with other GATA genes (termed exon Ia in Fig. 1A). We performed 5'-RACE analysis on RNA derived from human adult atria in order to determine precisely the transcriptional start site. The coding sequence spans at least six exons (II–VII). The existence of exon VII is deduced and is not indicated here, since we have yet to identify it in any of our genomic clones. The two initiator methionine codons within exon II are indicated. B, consensus splice sites flank exons II–VI. C, a sequence comparison of the human and mouse noncoding exons, Ia and Ib, is shown in uppercase type. 1 denotes the CAP site of exon Ia in each case. The A residue in the mouse exon, starred at position 127, denotes a minor downstream CAP site. 3'-Flanking genomic sequence, including splice donor sites, are in lowercase type. For Ib, the numbers indicated are genomic residues relative to the Ia CAP sites. The CAP site for human Ib has not been determined; thus, the genomic sequence that is adjacent to known cDNA sequence (uppercase type, not underlined) is shown in this alignment. 

GATA-6 Gene Organization, Expression, and Translation

FIG. 1. Structural organization of the human and mouse GATA-6 genes. A, the mammalian GATA-6 genes have two noncoding first exons (Ia and Ib), which both splice to a common acceptor site upstream of exon II. Exon Ib has an alternatively spliced 3'-end. The coding sequence spans at least six exons (II–VII). The existence of exon VII is deduced and is not indicated here, since we have yet to identify it in any of our genomic clones. The two initiator methionine codons within exon II are indicated. The sequence of this upstream exon showed extensive homology (90%) to a similar upstream region within the mouse GATA-6 locus. We therefore performed RT-PCR analysis on RNA isolated from both 10.5-day mouse embryos and adult mouse hearts, using a gene-specific primer from this region, together with a reverse primer complementary to exon II and demonstrated that a homologous untranslated 5' exon does exist within the mouse GATA-6 gene (termed exon Ia in Fig. 1). Again we determined the transcriptional initiation site by 5'-RACE performed on RNA derived from neonatal mouse hearts. The sequence of the major PCR product determined the position of the mouse exon Ia CAP site and demonstrated the size of the exon to be 243 bp (see Fig. 1C). A smaller, minor PCR product was also characterized from the 5'-RACE analysis,
we performed RT-PCR analyses on both adult and fetal human heart RNAs. In both cases, a single band was generated, the sequence of which confirmed that the human GATA-6 gene does indeed contain an alternatively expressed upstream exon, analogous to the mouse exon Ib. Alignment of the cDNA and genomic sequences revealed a splice-donor consensus sequence flanking the human Ib exon on the 3' side (Fig. 1C).

Thus, both human and mouse GATA-6 genes encode alternative noncoding upstream exons, transcribed from distinct promoter regions, approximately 850 bp apart in mouse, which could potentially be differentially regulated in a tissue- and/or developmental stage-specific manner. The 5' boundary of exon II is constant, regardless of which upstream exon is utilized. We have also found an alternatively spliced isoform of exon Ib, by RT-PCR analysis of 10.5-day total embryonic mouse RNA across the exon Ib-exon II splice junction. This RNA species is truncated by 170 bp at the 3' end of exon Ib, relative to that described above and again is spliced to the same acceptor site upstream of exon II (see Fig. 1). We assume this to be a minor species, since it was undetectable in RNase protection assays and was only apparent in RT-PCR analysis. We have not determined whether the human GATA-6 exon Ib is similarly subject to alternative splicing.

Transcripts Initiating from Both Noncoding Exons Are Widely and Cooperately Expressed in Fetal and Adult Mouse

We have determined the patterns of expression of the two mouse GATA-6 upstream exons (Ia and Ib) in RNA derived from both adult and embryonic tissues that express GATA-6 (including heart, duodenum, esophagus, and aorta). Exon Ia- or exon Ib-specific forward primers, together with a reverse primer complementary to exon II, were initially used in a qualitative RT-PCR analysis, and for all samples where GATA-6 was expressed, both exon Ia- and exon Ib-containing transcripts could be detected (data not shown). No evidence was ever found of transcripts containing both exon Ia and exon Ib sequences.

Quantitation of the GATA-6 transcripts was carried out by RNase protection analysis using probes spanning either the exon Ia/II boundary or the exon Ib/II boundary (Fig. 3A) on a range of tissues derived from different embryonic germ layers including duodenum (endodermal) and heart (mesodermal). Again, whenever GATA-6 could be detected, transcripts initiating from both noncoding exons were always found (Fig. 3B and data not shown), and exon Ia consistently appeared as the predominant GATA-6 RNA species. A detailed study was conducted of cardiac GATA-6 expression from embryonic day 11.5 mice through to adulthood. Once more, we found coordinate expression of exon Ia- and exon Ib-containing transcripts in the atria and ventricles at all stages analyzed (Fig. 3B).

During embryonic development, the atrial and ventricular levels of GATA-6 transcripts (including both Ia- and Ib-containing isoforms) were essentially equivalent, and they remained relatively constant between embryonic days 11.5 and 18 (Fig. 3C and data not shown). In the adult mouse heart, GATA-6 expression decreased, although atrial expression exceeded that of the ventricles. This preferential expression of GATA-6 within the atria was substantially amplified in the neonate; in 1-day pups, atrial GATA-6 expression was at least 2-fold higher than that seen in the ventricles (Fig. 3C), and this pattern was maintained for at least 14 days after birth (data not shown).

GATA-6 expression was also monitored by in situ hybridiza-
tion. Using a probe to detect all GATA-6 transcripts, we found a similar expression pattern to that reported previously (13) with expression in the heart, septum transversum, and umbilical vein of 9-day mouse embryos (Fig. 4A). By embryonic day 11, expression was additionally found in the urogenital ridge (Fig. 4B). The small signal in the head region was background

**Fig. 3.** RNase protection analysis of GATA-6 expression and 5′ noncoding exon utilization during mouse development. **A,** a schematic representation of the probes used for RNase protection analysis of total RNA extracted from mouse tissues. Antisense RNA probes were transcribed from cDNA templates comprising of 71 bp from the 3′ end of exon Ia and 221 bp from the 5′ end of exon II (mGATA-6XIa/XII) or 50 bp from the 3′ end of exon Ib and 221 bp from the 5′ end of exon II (mGATA-6XIb/XII). The expected sizes of probe fragments protected by transcripts initiating from exon Ia or exon Ib are indicated. **B,** a representative autoradiograph of RNase protection analysis carried out on total RNA from pooled mouse tissues (from 2–26 mice or embryos) at various stages of development. The sizes of the protected probe fragments (indicated by arrows) correspond to those predicted in A. **C,** mean expression (of two determinations) of total GATA-6 was quantified by PhosphorImager analysis and normalized to the glyceraldehyde-3-phosphate dehydrogenase loading control. tRNA denotes the negative control lane containing yeast tRNA.
staining that was also observed in the sense control (data not shown). Within the heart, expression of GATA-6 was strongest in the endocardial cushions of the atrioventricular canal and outflow tract (arrow in Fig. 4C). Later on, when the elongation had arrested, GATA-6 levels were low within the septum itself but high within the cells that were still undergoing morphogenetic changes within this region (Fig. 4D). Thus, relatively high levels of GATA-6 were detected within a very dynamic population of endocardial cushion and septal precursor cells. This association is seen in both mouse and humans (data not shown) and may be related to the increased expression of GATA-6 within the early postnatal atria (Fig. 3C), a period when the interatrial septum is undergoing substantial remodeling (Ref. 28; also see “Discussion”).

In an attempt to determine whether there was heterogeneous expression of the noncoding GATA-6 exons, we used exon-specific probes for in situ hybridization. Using a mouse exon Ib-specific riboprobe, we found expression profiles identical to those observed using riboprobes to detect all mouse GATA-6 RNA species (Fig. 4, compare D–F with A–C). The use of exon Ia-specific riboprobes was unsuccessful in in situ hybridization analyses. We attribute this to the small size and high GC content of the probe. This meant that there was little incorporation of Dig-UTP and also that there was a high level of RNA secondary structure that could have impeded specific hybridization. This combination led to a very low specific signal, which was difficult to distinguish from the background level of staining. Thus, while in situ hybridization could not elaborate on the RT-PCR and RNase protection information about the expression pattern of exon Ia, it has shown that exon Ib is

---

**Fig. 4. Expression of GATA-6 in vivo.** In situ hybridizations, using antisense mRNA probes specific to either all GATA-6 transcripts (A–C) or exon Ib only (D–F) were performed on serial sections of 9-dpc (transverse sections; A and D) and 11-dpc (sagittal sections; B, C, E, and F) mouse embryos. GATA-6 is present in the heart, septum transversum, and umbilical vein in the 9-dpc embryo. Later, in the 11-dpc embryo, GATA-6 is again detected in the heart and umbilical vein and is also present in the urogenital ridge. Levels in the heart are strongest in the endocardial cushions of the outflow tract and the atrioventricular canal as well as the forming interventricular septum (arrow in C). In all cases, the pattern of the staining derived from the exon Ib probe matched exactly that of the pan-GATA-6 probe. It can also be seen that in the mouse 12.5-dpc heart (G and H) GATA-6 is expressed most strongly in the area undergoing morphogenesis, opposite the interventricular septum, which will eventually contribute to the final separation of the ventricles (arrow in H). A, atria; EC, endocardial cushions; H, head; Ht, heart; IVS, interventricular septum; NT, neural tube; OT, outflow tract; SP, septum primum; ST, septum transversum; UR, urogenital ridge; UV, umbilical vein; V, ventricle.
expressed in all GATA-6-positive tissues at the time points analyzed.

**Sequence Analysis of the Promoters and Comparisons between Species**

As far as we have been able to establish, the human and mouse GATA-6 genes show similar patterns of tissue-specific expression, particularly in the heart region during cardiogenesis. Therefore, similar molecular mechanisms may regulate their tissue-specific expression. Moreover, DNA sequences within the GATA-6 loci that modulate this regulation will be evolutionarily conserved between the mammalian species. We therefore sequenced both the mouse and human GATA-6 loci from approximately 1 kb upstream of the CAP site of exon Ia through to exon II (EMBL accession numbers AJ243146 and AJ245649, respectively).

The immediate 5'-flanking regions of exon Ia show extensive homology (73%) over the most proximal 600 bp, suggesting a regulatory function. Upstream of this, homology is not apparent. Similarly, the region of approximately 600 bp between the exons Ia and Ib of mouse shows great similarity (approximately 80%) to the analogous region in humans. We were surprised to note, however, that the most striking evolutionary conservation of 5' noncoding genomic DNA sequence between the mammalian GATA-6 loci has occurred within exon Ia. Thus, the homology between the human and mouse (noncoding) exon Ia sequences is >90% over their entire lengths (Fig. 1C). This is greater than the homology at the nucleotide level of the coding portions of the mouse and human GATA-6 cDNAs (approximately 85%). By contrast, the homology between the mouse exon Ib and the corresponding region in human was not nearly so striking (Fig. 1c). In addition, the intron regions of the two mammalian species between exon Ib and exon II are not apparently similar.

We also compared the DNA sequence of the mammalian GATA-6 genes with the 5' region of the chicken GATA-6 locus in order to determine whether there were putative regulatory regions conserved between more evolutionarily diverse species. Unlike the mammalian genes described here, the chicken GATA-6 gene has a single untranslated upstream exon, although its pattern of expression is similar, particularly in cardiac cells. We therefore noted with great interest that the only region of obvious homology between the chicken, mouse, and human GATA-6 genes (other than within coding sequences) lay between a region within exon Ia of the mammalian genes and a region of the chicken promoter, upstream of its (single) untranscribed exon (see Fig. 5A). Thus, it seems likely that sequences that act to regulate the rate of transcription initiation reside within a transcribed region (exon Ia) of the mammalian GATA-6 genes. By contrast, the sequences of the chicken untranslated 5' exon showed no apparent similarity to any region of the mammalian GATA-6 loci.

---

**Fig. 5.** Sequences within the non-coding exon Ia can act to promote GATA-6 transcription in a transient transfection assay. A, a region within the chicken GATA-6 core promoter, upstream of the (single) CAP site (9), shows good homology to sequences within the mouse and human noncoding Ia exon. The numbers indicated are relative to the CAP sites of exon Ia for mouse and human or the single transcription initiation site for chicken. B, mouse genomic fragments that span the putative promoter regions of mouse GATA-6 were assayed for their abilities to direct CAT reporter gene expression in WRL cells; the CAT data from duplicate transfection assays were normalized to expression from the co-transfected SV-40-driven β-galactosidase internal control. The transcripts resulting from the three positive constructs are indicated; in each case, transcripts initiate from multiple sites within exon Ia and are spliced from the donor site 3' of exon Ia to the acceptor site within the CAT transcription unit.
Sequences within Exon Ia Act to Promote Transcription in Transient Transfections

We tested the ability of the mouse exon Ia sequences to stimulate transcription by transient transfection of CAT reporter constructs in a human fetal liver cell line, WRL, which we have shown by RNase protection assays expresses GATA-6. Moreover, by RT-PCR analysis we have demonstrated that both Ia- and Ib-containing isoforms of the GATA-6 transcript are expressed in these cells (data not shown). We constructed plasmids where CAT expression was driven by either a 942-bp Kpn–NheI fragment, comprising 906 bp of sequence immediately 5' to the in vivo Ia CAP site in addition to 36 bp of exon Ia (MXIa-CAT), or a 1149-bp Kpn–XhoI fragment that comprised the same upstream sequences in addition to the entire length of exon Ia (MXIa + NX-CAT; see Fig. 5B). Constructs were co-transfected with pSV-lgal (Promega) to act as an internal control, and resulting CAT levels were compared with those resulting from the transfection of the promoterless pCAT-Basic. As shown in Fig. 5B, the upstream sequences alone, together with a minimal 5’ region of exon Ia, were insufficient to drive any CAT expression in these cells above that resulting from pCAT-Basic. However, inclusion of the entire exon Ia resulted in a dramatic stimulation of CAT activity of some 40-fold, suggesting that transcribed sequences within the mouse GATA-6 gene act to promote transcription. We also tested the ability of the NheI–XhoI fragment (NX-CAT), comprising the whole of exon Ia apart from the most 5’ 36 bp, to function as a promoter in this assay in the absence of the in vivo CAP site. As shown in Fig. 5B, expression from this construct was approximately 30-fold higher than that resulting from pCAT-Basic, suggesting that this fragment constitutes a bona fide promoter, at least in the context of this transfection assay. Mapping the sites of transcription initiation resulting from the MXIa + NX-CAT and NX-CAT transfections by 5'-RACE analysis found initiations from multiple sites within the NheI–XhoI fragment in both cases. These data suggest that transcribed sequences within exon Ia act positively to regulate the initiation of transcription of mouse GATA-6 but that the precise site of initiation in vivo may be influenced by more distal sequences and/or tissue-specific factors.

As stated previously, the WRL cell line also expresses endogenous GATA-6 transcripts containing the Ib exon, suggesting that regulatory factors necessary to promote transcription from the exon Ib CAP site are present in these cells. Given the high degree of extended homology observed in the sequences between exon Ia and Ib, we tested the ability of this region to promote transcription of the Ib exon. A PCR-generated genomic fragment that comprises all of the sequence between exons Ia and Ib in mouse in addition to most of exon Ib (MXIb-CAT) was used to drive CAT expression in our transient assay. As shown in Fig. 5B, only minimal CAT activity (less than 2-fold above the pCAT-Basic background level) resulted from these transfections, suggesting that additional, more distal sequences are necessary for the expression of Ib. Since the expression patterns of the Ia and Ib exons are so often quantitatively and qualitatively so similar, we wondered whether a single regulatory element might coordinate transcription initiation from both CAP sites. We therefore extended the 5’ boundary of MXIb-CAT to include the NheI–XhoI fragment within exon Ia, which we have shown to affect transcription from exon Ia. This construct (MXIb + NX-CAT) directed the highest levels of CAT activity of all our constructs in WRL cells (Fig. 5B). However, by 5’-RACE analysis we demonstrated that all of the CAT transcripts had initiated from within exon Ia (i.e. within the NheI–XhoI fragment) and had become spliced from the donor site 3’ of exon Ia, to the acceptor site within the CAT transcription unit. We detected no transcripts that contained sequences from exon Ib. This suggests that more distal sequences (either 5’ or 3’) are required to positively regulate transcription of exon Ib.

Translation of GATA-6 in Humans and Mice Initiates at Alternative Initiator Methionine Codons

The open reading frame (ORF) of the human GATA-6 cDNA clone has been reported to encode a protein of 449 amino acids extending from an initiator methionine codon at 716 bp (14, 15), which aligns at its N terminus to other members of the GATA-4, -5, and -6 subfamily, reported in various species (4, 7, 13, 29). However, analysis of our cDNA clone has revealed a longer potential ORF encoding a protein of 595 amino acids, which commences at a more upstream, “in-frame” putative initiator methionine codon at 278 bp (14). Both the methionine codons at 278 and 716 bp are within a theoretically favorable context for translation initiation (30–32) and are located within exon II. An in-frame stop codon is located 33 bp upstream of the more 5’ AUG (at 245 bp), also within exon II. Additionally, analysis of the nucleotide sequence we have generated from mouse GATA-6 cDNA and genomic clones revealed a similar potential 5’-extended ORF in the mouse gene, in addition to that previously reported (13). As in the human gene, both potential methionine initiator codons are located within exon II (see Fig. 1A). However, the upstream in-frame stop codons within the mouse gene are located within exons Ia or Ib (138 and 48 nucleotides upstream, respectively). An alignment of the putative peptide translations of the ORFs of the human and mouse GATA-6 genes reveals a high level of homology throughout the entire sequence of the proteins, including the extended N-terminal region described here (Fig. 6A).

Translation from the more upstream methionine codon would generate a polypeptide sequence commencing with MALT in both of the mammalian genes, and that from the more downstream AUG would result in the N-terminal motif MYQ, as reported previously (13–15). The extended N-terminal sequence comprises 146 and 147 amino acid residues in humans and mice, respectively. To test the ability of the AUG codons to initiate translation, we programmed a coupled transcription and translation system (TNT, Promega), in the presence of [35S]methionine, with either a human GATA-6 cDNA construct, which contained the full predicted coding sequence (MALT-ORF; see Fig. 6B), a construct truncated upstream of the published AUG (MYQ-ORF), or a construct identical to MALT-ORF except that AUG at codon 147 of the ORF had been mutated to a leucine codon (M147L; Fig. 6B). Leucine was chosen because of its poor initiation potential and because the replacement of methionine with leucine is not anticipated to change the functional properties of the mutated protein (33). All of the constructs included upstream cDNA sequence corresponding to the Kozak consensus. Two translation products corresponding to 64 and 52 kDa were detected in the MALT-ORF translation (Fig. 6B). Translations programmed with the MYQ-ORF construct showed only the smaller, 52-kDa band, while that programmed with the M147L construct showed only the 64-kDa species (Fig. 6B). We confirmed that the two bands represent 52- and 64-kDa isoforms of human GATA-6 by Western blot analysis using a GATA-6-specific antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) (Fig. 6B). Translation of the mouse full-length construct also gave two bands of 50 and 59 kDa (data not shown). Thus, in vitro, translation can initiate from two alternative initiator methionine codons within the mammalian GATA-6 genes, giving rise to two protein isoforms.

A. Davies, unpublished results.
transfection of the GATA-6 expression construct with a luciferase gene comprising mutated GATA-binding sites in its promoter demonstrated that the increase in expression was mediated through the GATA sites (Fig. 7). The transactivation activity of the full-length mouse GATA-6 expression construct was similar to that of the human MALT-ORF construct (data not shown).

To determine the transactivation potential of the two human GATA-6 isoforms separately, we further compared the luciferase activity resultant upon transfection of M147L and MYQ-ORF. Levels of expression directed by M147L, which produces only the longer polypeptide, were at least as high as those directed by MALT-ORF (which makes both forms); however, deletion of the N-terminal 146 amino acids (MYQ-ORF) reduced luciferase expression by 50% (Fig. 7). Thus, the extended N-terminal portion of GATA-6 contains transactivation activity, apparent in this assay.

The construct MALT-146, in which these N-terminal 146 amino acids were linked directly to the GATA-6 binding domain (Fig. 7), was unable to significantly stimulate target gene expression above levels resultant upon either empty vector alone (pCS-2+) or the DNA-binding domain linked to a minimal N-terminal sequence (MALT-18; Fig. 7). Two activation domains have previously been reported within the mouse GATA-6 gene in the region homologous to that deleted in MALT-146 (34). Thus the novel N-terminal activation region described here may require interaction with one or both of these previously described domains, or to be spaced away from the (DNA-binding) zinc fingers, in order to display its transactivation potential.

We confirmed the existence of these transactivation domains in human GATA-6 by comparing luciferase activity resultant from co-transfection of a deletion series of MYQ-ORF. Thus, the transactivation activity of the MYQ-ORF construct was reduced 3.5-fold by deletion of a domain proximal to the zinc finger region (MYQ-77) and reduced a further 3.5-fold by the further deletion of a more distal domain (MYQ-20; Fig. 7). EMSAs performed upon transfected cell extracts with an oligonucleotide encoding the core WGATAR motif from the mouse a1 globin gene promoter (aG2; Ref. 35) confirmed that equivalent amounts of GATA-6 protein were generated in these experiments (data not shown).

We were unable to demonstrate transactivation activity with the GATA-6 derivatives in transfected COS cells, which do not have detectable endogenous GATA-6 expression. However, by EMSAs we could show that GATA-binding activity did accumulate in the transfected cells (data not shown). This suggests that GATA-6 may require co-factors that are not present in COS cells to function as a transcriptional activator.

The 5'-UTR Does Not Affect the Rate or Site of Translational Initiations in Vitro

In addition to providing a mechanism for the complex differential regulation of transcription, alternative promoter usage results in the differential expression of 5'-untranslated regions, which can influence the translation of a gene (for a review, see Ref. 36). Thus, the translational efficiency and/or alternative methionine usage of the two GATA-6 mRNA isoforms with different leader exons may vary. We therefore tested the ability of two constructs, in which either mouse exon 1a or exon 1b was linked to exon II, to direct translation in a cell-free transcription-translation coupled system (TNT; Promega). As shown in Fig. 8, equivalent levels of both polypeptides were detected from the different cDNA isoforms, with the ratio of the two bands resulting from the alternative use of the two in-frame initiator codons the same in each case. Thus, in
vitro, the differential usage of 5'-untranslated sequence does not appear to regulate the rate of translation initiation or the choice of initiator methionine.

**DISCUSSION**

We demonstrate here that the human and mouse GATA-6 genes are transcribed from two distinct promoter regions, resulting in exon Ia- and exon Ib-containing isoforms of the GATA-6 transcript. The use of alternative promoters and transcriptional start sites is a common mechanism employed for the differential regulation of a gene that displays a complex temporal and/or spatial expression pattern (37). Each member of the GATA family of transcription factors shows just such a complex pattern of tissue-restricted and developmental stage-specific expression (4, 29, 38, 39). It is not surprising therefore that there are now several reports of distinct promoters mediating the cell type-specific transcriptional regulation of GATA genes in different species. Thus, in the case of the mouse GATA-1 gene, a distal promoter gives rise to testis-specific expression, while a more proximal promoter directs transcription in hematopoietic cells (2, 40, 41). Similarly, an upstream promoter within the mouse GATA-2 gene is active only in hematopoietic cells, while a downstream region directs GATA-2 transcription in all cells in which it is expressed (10). Two distinct transcriptional start sites have also been reported for the chicken GATA-2 gene, although the tissue distribution of promoter usage differs from the mouse (9, 11). Most notably, the predominant chicken GATA-2 transcript within erythroid cells derives from the proximal promoter, demonstrating evolutionary diversity in the regulation of the GATA-2 gene between avian and mammalian species. The chicken GATA-5 gene also has two promoters that appear differentially regulated in the heart in a developmental stage-specific manner (12).

In the case of the chicken GATA-6 gene, the expression of only one 5'-noncoding exon has been reported, which further indicates the evolution of regulatory mechanisms that may be divergent between species (9). It is perhaps surprising, however, that we detected both exon Ia- and exon Ib-containing isoforms of the GATA-6 transcript by RNase protection analyses in all tissues in which GATA-6 is expressed. For example, cells of mesodermal (cardiac) and endodermal (duodenal) origin both co-express the two RNA isoforms. Moreover, wherever we have looked, the relative levels of the two transcripts remained roughly constant. Thus, exon Ia expression was consistently slightly higher than that of exon Ib (less than 2-fold) within both different cell types and within cardiac cells at different developmental stages (data not shown and Fig. 3B). Clearly, we...
cannot preclude that subtle changes in the promoter usage do occur, perhaps in very specific subsets of cells and/or at particular developmental stages. An example of subtle developmental stage-specific regulation of alternative promoters has been demonstrated in the case of the p45 subunit of the human NF E2 (42). As in the case of the GATA-6 gene described here, two forms of the NF E2 gene transcript that are transcribed from distinct promoters give rise to the same protein and are co-expressed in both fetal liver and adult bone marrow. However, the two isoforms are expressed in different ratios during development. The independent regulation of two promoters that are both used within the same cell type may provide the flexibility necessary to “fine tune” the absolute cellular levels of the GATA-6 protein, which we have shown to be critical for correct cardiac differentiation (7).

Alternative promoter usage, in addition to providing a mechanism for differential transcriptional regulation, gives rise to transcripts with different 5′-UTRs. The 5′ noncoding regions of a number of cellular mRNAs encoding regulatory proteins have been shown to provide post-transcriptional regulation during embryonic development (for a review, see Ref. 36). The precise mechanisms involved in 5′-UTR-mediated control are not well understood; however, mRNA secondary structure and/or the binding of specific transacting factors have been implicated. We therefore analyzed the influence of the alternative 5′ noncoding regions upon translation of the mRNAs in vitro. We found that both Ia- and Ib-containing mRNA species were translated equally efficiently in our assay. Moreover, the ratio of the two protein species that resulted from the alternative use of the two translation initiator codons was the same in each case. Thus, we were unable to detect a differential effect of the alternative 5′ exons upon translation in vitro; we cannot, of course, rule out an effect in vivo.

There is a striking degree of homology (90%) between human and mouse sequences within the more upstream (Ia) noncoding exons of the GATA-6 genes. This is much greater than the homology between noncoding exons of most other mammalian genes (or GATA-6 exon Ib), suggesting that this region may have an evolutionarily conserved regulatory function. Although the promoters of cellular genes are classically assumed to be 5′ to the transcriptional initiation site, there are several examples of regulatory DNA motifs, which bind transcription factors, within the transcription unit itself (43–46). Moreover, sequences within noncoding exons have in a number of cases been shown to affect the rate of transcription initiation (47–50). We tested the function of the mouse exon Ia sequences by transient transfections into GATA-6-expressing cells and found them to act as a classical promoter, even in the absence of the in vivo CAP site or any sequences upstream. When the sequences of the mouse and chicken GATA-6 loci upstream of exon II were aligned, the only regions displaying significant homology were sequences within exon Ia of the mouse gene and a region within the (single) chicken GATA-6 core promoter, upstream of the reported CAP site (9). It is possible, therefore, that in the case of the mammalian genes, mechanisms have evolved to direct transcriptional initiation from upstream of an ancestral promoter region, which again suggests subtle evolutionary differences between avian and mammalian species in the regulation of gene expression. In our expression assay, however, the mouse exon Ia sequences acted to direct the initiation of transcription, albeit in a heterogeneous fashion, from downstream of the normal in vivo CAP site. This may represent an ancestral mechanism, which can occur in the absence of more distal sequences and/or specific factors that are required for correct initiations. The minor CAP site that we detected arising from within mouse exon Ia (see Fig. 1C) may indicate that this “ancestral promoter” functions in vivo. It is intriguing that human and mouse GATA-3 5′ noncoding exons show a similarly high degree of homology (51, 52) and that the mouse GATA-3 gene has a tissue-restricted DNase I-hypersensitive site that maps to within this exon, further suggesting a regulatory role in GATA-3 expression (53).

We also tested the ability of the regulatory exon Ia sequences to enhance transcription from the downstream CAP site (Ib). In the case of the bovine growth hormone receptor gene, ubiquitous expression from two promoters is co-regulated, and it has been reported that one promoter can serve as an enhancer for the other (54). However, although we did observe strong expression from the MXIb + NX-CAT construct (see Fig. 3), all of the transcripts initiated from within exon Ia. This suggests that the expression patterns of Ia- and Ib-containing GATA-6 RNA isoforms are independently regulated. Moreover, this experiment demonstrates that the minimal promoter region for the GATA-6 Ib exon is not contained within the sequences of MXIb + NX-CAT. As mentioned above, the differential usage of two 5′ exons is not a feature of the regulation of chicken GATA-6 (9). However, a positively acting regulatory region has been detected within the first intron of the chicken GATA-6 gene (9); perhaps, therefore, this region has evolved to direct exon Ib transcriptional initiation in mammals.

By RT-PCR, we have also detected in 10.5-day embryos an alternatively spliced variant of the exon Ib-containing isoform of mouse GATA-6 (Ib′). Alternative mRNA splicing, giving rise to different protein isoforms, is a feature of several GATA family members including Xenopus and chicken GATA-2 (11, 55, 56) and chicken GATA-5 (12). However, in the case of the alternative splice described here, a common splice acceptor site upstream of the translational start sites in exon II is utilized; thus, the ORF of the protein is unaffected. The Ib′ RNA species was not detected in our 5′-RACE analysis; thus, we do not know where these transcripts are initiated, and we have not as yet determined the pattern of expression of this splice variant. Nonetheless, it is clear that the expression of the mammalian GATA-6 gene is subject to complex regulatory mechanisms at the level of transcription initiation and processing.

We demonstrate here that the human and mouse GATA-6 genes each encode a longer polypeptide, in addition to the one reported previously (13–15), which results from the initiation of translation at a more upstream methionine codon. Moreover, in an in vitro translation assay, we detect two protein isoforms, resultant upon translation initiation from the two potential initiator methionines within our cDNA clones. The longer protein appeared more abundant in this assay, suggesting that initiations from the more upstream methionine are favored, at least in vitro (see Fig. 6). We have attempted to determine the methionine usage in vivo by Western blotting using GATA-6-specific antibodies (Santa Cruz Biotechnology and a kind gift of M. Maeda). However, as we have reported previously, these antisera fail to detect in vivo GATA-6 levels by immunoblotting procedures (5).

We have transfected GATA-6 constructs that direct expression of the two protein isoforms into HepG2 cells. We show that the protein with the extended N-terminal region has a higher potential to transactivate expression of a reporter gene driven by multimerized GATA-binding sites than that resulting from translation initiation at the downstream methionine. Alternative translation initiator sites have also been reported for the mouse GATA-1 gene (57). In this case, two protein isoforms are expressed in vivo, which differ by the presence or absence of 83 amino acids at their N termini. The isoforms differ in their transactivation potentials, and their expression in hematopoietic cells appears developmentally regulated. The amino acid
sequences of the GATA-1, -2, and -3 subfamily from their N termini (1), and the alternative methionine within the mammalian GATA-1 gene described (57) results in translation initiation internal to this alignment. The previously reported amino acid sequences of the GATA-4, -5, and -6 subfamily also show great homology at their N termini, and all polypeptides begin with the MYQ peptide motif (29). Thus, the alternative methionine usage described here, upstream of the MYQ motif, suggests divergent evolution within the GATA-4/5/6 subfamily.

We do not yet know whether this N-terminal extension represents an evolutionarily conserved feature of GATA-6 genes or is specific to mammalian genes; however, we note that potential 5′-extended ORFs exist within both the zebrafish and chicken GATA-6 genes.⁶

In addition to sequence homology, members of the GATA-4/5/6 subfamily also manifest overlapping expression domains in heart and endoderm and, in some cases, appear to be able to compensate for each other’s loss (58). However, the precise functional role of these factors remains to be determined. We have previously suggested that the early expression of GATA-6 in precardiac mesoderm may play a role in keeping such cells in an undifferentiated state (7). We have now shown, consistent with previous studies, that GATA-6 is also expressed later in embryogenesis in both atrial and ventricular compartments of the developing mouse heart and that the postnatal expression is higher in the atria. We have further demonstrated that the endocardial cushions are early sites of relatively high GATA-6 expression. These cushions will later form the valves and septa of the heart (59), and it is of interest that the early expression of GATA-6 with cardiac cells associ-...