Mammalian Genome 10, 993–999 (1999).

Mammalian Genome
Incorporating Mouse Genome

Functional analysis and chromosomal mapping of *Gata5*, a gene encoding a zinc finger DNA-binding protein

Georges Nemer,[1] Salman T. Qureshi,[2,3] Danielle Malo,[2,3] Mona Nemer[1]

[1]Laboratoire de développement et différenciation cardiaques, Institut de recherches cliniques de Montréal, 110 des Pins Ouest, Montréal, Québec, Canada H2W 1R7 and Département de Pharmacologie, Université de Montréal, Montréal, Québec, Canada
[2]Department of Medicine and Human Genetics, McGill University, Montréal, Québec, Canada H3G 1Y6
[3]Centre for the Study of Host Resistance, L11-144 Montréal General Hospital, 1650 Cedar, Montréal, Québec, Canada H3G 1A4

Received: 17 December 1998 / Accepted: 9 June 1999

Abstract. The GATA family of zinc finger proteins are transcriptional regulators with critical functions in lineage differentiation and embryonic development. Based on structural and expression pattern comparisons, the GATA proteins have been subdivided into two groups. The first subgroup consists of GATA-1, -2, and -3, which are all highly expressed in the hematopoietic system, whereas GATA-4, -5, and -6 are present essentially in the heart and gut. We have isolated and functionally characterized the rat GATA-5 cDNA, which encodes a 45-kDa protein with 71%, 73%, and 97% homology to its amphibian, avian, and murine homologs, respectively. Northern blot analysis showed that rat GATA-5 is expressed in a dynamic pattern during embryonic and postnatal development. In the midgestation embryo, GATA-5 transcripts are most abundant in the heart and decrease dramatically in the postnatal heart; in contrast, GATA-5 expression is upregulated in the lung and gut during postnatal development. Functional studies with recombinant GATA-4, -5, and -6 proteins show that GATA-5 has preferential affinity for a subset of GATA elements found on cardiac promoters and differentially activate cardiac gene transcription. Structure-function analysis revealed the presence of an activation domain within the carboxy terminal region of GATA-5 that is essential for transcriptional regulation of target promoters. Linkage analysis localized *Gata5* to distal mouse Chromosome (Chr) 2 in a conserved linkage group with genes localized to rat Chr 3q43 and human Chr 20q13.2-q13.3. The results suggest that GATA-5 may have specific downstream targets and that GATA-4, -5, and -6 can only partially substitute for each other in cardiogenesis. Thus, *Gata5* probably plays a specialized evolutionary conserved role in cardiac development.

Introduction

GATA proteins are lineage-restricted zinc finger transcription factors that play critical roles during cellular growth and differentiation. Indeed, the three GATA factors that are expressed in hematopoietic derived cells are required for embryonic development and have nonredundant essential functions for normal hematopoiesis (Pandolfi et al. 1995; Tsai et al. 1994). GATA-1, the founding member of the family, is expressed in all erythroid precursors, and targeted disruption of its gene arrests erythropoiesis at the proerythroblast stage both in vivo and in vitro (Pevny et al. 1995; Weiss et al. 1994). Similarly, inactivation of the GATA-2 gene, whose expression is very high in hematopoietic stem cells, interferes with normal hematopoiesis by inhibiting expansion of hematopoietic progenitor cells (Tsai et al. 1994). Hematopoiesis is also impaired in the absence of GATA-3 (Pandolfi et al. 1995).

Correspondence to: M. Nemer

Three other GATA factors, GATA-4, -5, and -6, were recently identified (Grépin et al. 1994; Jiang and Evans 1996; Kelley et al. 1993; Laverriere et al. 1994). Based on sequence homologies and tissue distribution—mainly in the heart and gut—it has been suggested that GATA-4, -5, and -6 form a subfamily of GATA factors that may have important roles during cardiogenesis. The expression pattern of all three genes in the developing heart and the functional analysis of GATA-4 support this hypothesis. Thus, both in amphibians and in mammals, GATA-4 transcripts are detected in precardiac cells and persist throughout the embryonic and postnatal heart (Grépin et al. 1994; Heikinheimo et al. 1994; Jiang and Evans 1996). In Xenopus, axis disruption experiments show that GATA-4 transcription correlates with specification of cardiac progenitors (Jiang and Evans 1996), and in adults, GATA-4 expression is restricted to the heart and gonads (Grépin et al. 1994; Laverriere et al. 1994). Consistent with a role for GATA-4 in myocardial differentiation, cardiac-specific transcription of several cardiac marker genes requires conserved GATA binding sites, and recombinant GATA-4 is a potent transactivator of their promoters (Grépin et al. 1994; Ip et al. 1994; Molkentin et al. 1994). In an in vitro stem cell model, GATA-4 transcription correlates spatially and temporally with cardiac differentiation, and inhibition of GATA-4 expression blocks formation of terminally differentiated beating myocytes, while ectopic expression of GATA-4 potentiates cardiogenesis, suggesting a role for GATA-4 in development of the cardiogenic lineage (Grépin et al. 1995, 1997). Consistent with this, inactivation of the *Gata4* gene in mice resulted in embryonic lethality owing to failure of the embryos to form a primitive heart tube (Molkentin et al. 1997).

The expression pattern of GATA-4 in the myocardium of the developing heart overlaps with that of GATA-6, although it remains uncertain whether both factors colocalize or whether they are differentially expressed in distinct cell populations (Jiang and Evans 1996; Tamura et al. 1993). GATA-6 is also more broadly expressed than GATA-4, most notably in stomach and small intestine of birds, amphibians, and mammals (Jiang and Evans 1996; Laverriere et al. 1994; Tamura et al. 1993). While the *Gata5* gene is also transcribed in the precardiac mesoderm of birds and amphibians (Jiang and Evans 1996; Laverriere et al. 1994), its expression pattern is different from that of *Gata4* and *6* as it is less abundant in the dorsal precardiac lateral plate and is detected in more posterior regions of the ventral midline (Jiang and Evans 1996). Moreover, unlike GATA-4 and GATA-6, expression of GATA-5 is found predominantly in the endocardium of the embryonic heart of both species, raising the possibility that GATA-5 might play a specific role in endocardial development.

We report the cloning of the rat GATA-5 cDNA, its functional characterization, and the chromosomal localization of *Gata5* in the mouse genome. The results indicate that the *Gata5* gene has been

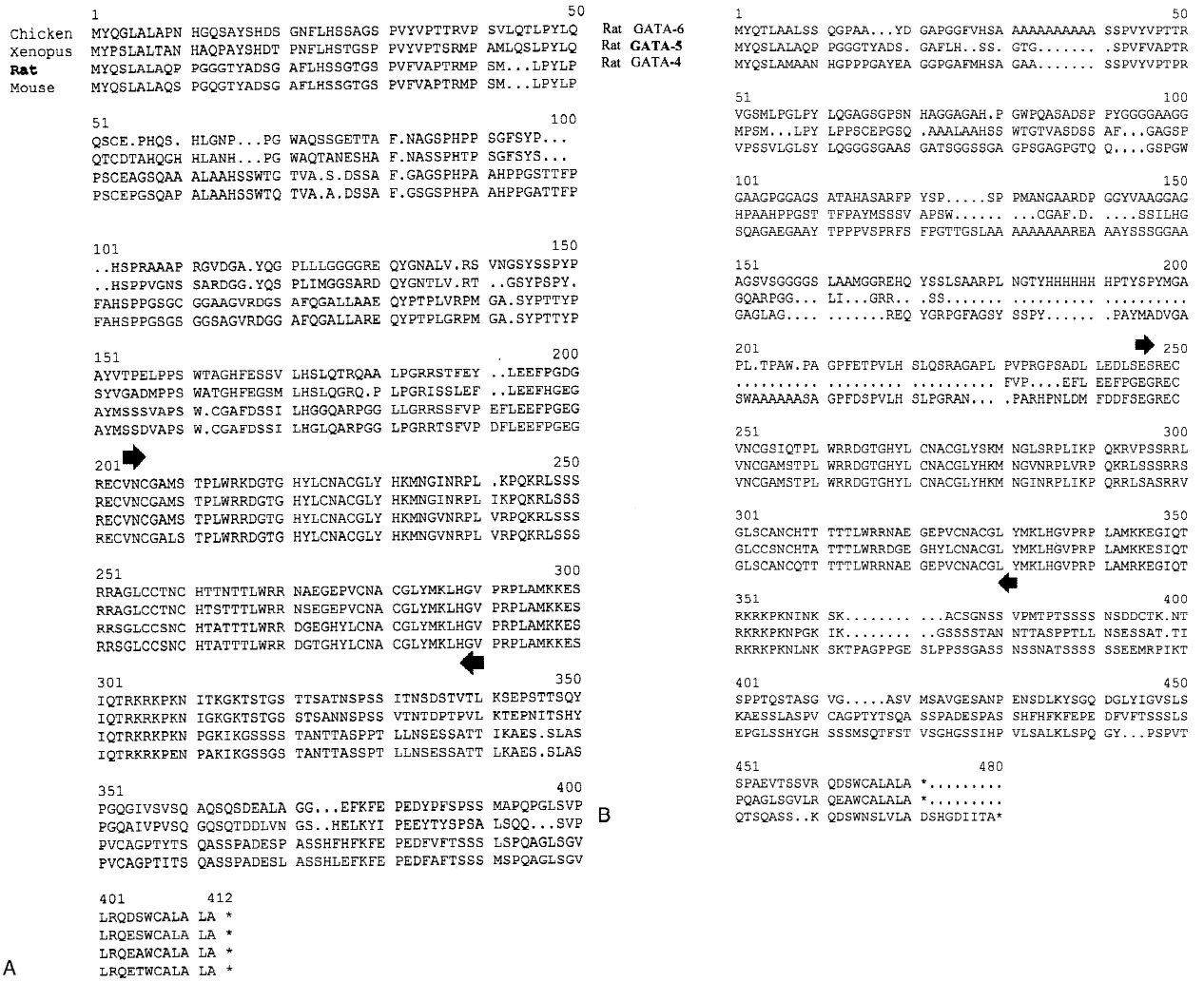

Fig. 1. Analysis of rGATA-5 primary sequence. **(A)** Comparison between the mouse, chicken, Xenopus, and rat GATA-5 proteins. The computer-assisted sequence analysis shows that the overall homology is about 72% between the rat and Xenopus and chicken, and 94% between the rat and mouse. The zinc finger domains delimited by the arrows are >94% homologous. **(B)** Amino acid sequence comparison between the rat GATA-4 (access no. L22761), GATA-5 (access no., submitted), and GATA-6 (access no. L22760). The homology is high in the zinc finger DNA-binding domains (>90%), whereas the overall sequence homology varies between 58% and 65%.

evolutionarily conserved from amphibians to mammals and that GATA-5 has higher affinity than the other cardiac GATA proteins to some GATA elements; consequently, some GATA-dependent cardiac promoters—such as endothelin 1 (ET-1)—are preferential GATA-5 targets. These results suggest a distinct role for GATA-5 in differentiation of the cardiogenic lineage.

Materials and methods

Cloning and DNA sequencing. Total RNA was extracted by the guanidinium–thiocyanate–phenol/chloroform method (Grépin et al. 1994) from hearts of 3-day-old Sprague Dawley rats. cDNA was synthesized from 5 μg total RNA with dT_{12-18} adaptor primer and the avian myeloblastosis virus reverse transcriptase enzyme(AMV-RT). The cDNA was amplified with two degenerate oligonucleotides in the sense: TA(CT) CA(CT) AA(AG) ATG AA(CT) GG and antisense orientation: CG(AG) TG(AGCT) A(AG)(CT) TTC AT(AH) TA based on the sequence of Xenopus and chicken GATA-5, which amplifies a 177-bp fragment. The PCR product was cloned in Bluescript vector (pKS) /*Eco*RV and sequenced with universal and reverse primers. Two specific sense and antisense primers were synthesized according to the sequence read, and PCR was carried in the same manner as above with, on the one hand, a sense primer and an antisense oligonucleotide corresponding to the region spanning the stop codon (TAGGCCAGGGCCAGGGCACACCAG), and, on the other hand, a sense oligonucleotide corresponding to the region around the start codon (ATGTATCAGGGTCTGGCTC), and a specific GATA-5 antisence primer. The two fragments obtained were cloned in pKS/*Eco*RV and sequenced.

Northern blots and RT-PCR. Total RNAs from different embryonic, neonates, and adult tissues of Sprague Dawley rats were extracted by the guanidinium-thiocyanate method. Samples of 20 μg were loaded on formaldehyde agarose gels, transferred to Nytran-nylon membranes (Schleicher & Schuell), and hybridized to a random primed radiolabeled ^{32}P-rGATA-5 probe corresponding to the entire coding region. Blots were washed twice for 15 min at 65°C in 0.1 × SSC, 0.1% SDS, and exposed for 4 days at −70°C. The membranes were then stripped and reprobed with the 18S probe as previously described (Grépin et al., 1994).

For PCR amplification, two oligonucleotides in the sense: GAACCAGTGTGCAACGCCTG and antisense orientation: CTAGGCCAGGGCCAGGGCACACCAG were used on cDNA transcribed from the RNA extracted from different tissues. The PCR products corresponding to the C-terminal region of GATA-5 (nucleotides 806–1215) were separated by agarose gel electrophoresis, transferred to nylon membranes, and hybridized with the randomly radiolabeled GATA-5 DNA probe (corresponding to the entire coding region), washed for 15 min at 65°C with 0.1 × SSC, 0.1% SDS, and exposed with Kodak Xomat/AR film for 1 h at room

temperature. The same cDNA samples were subjected to amplification of rat tubulin transcripts as previously described (Grépin et al. 1997).

Cell culture and transfections. HeLa and CV1 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum. DNA was transfected by the calcium phosphate precipitation technique 24 h after plating. The amount of the reporter expression vector was kept constant at 3 μg per 35-mm culture dish. The cells were harvested 36 h after transfection, and luciferase activity was measured by an LKB luminometer. The results are the mean of three independent experiments done each in duplicate. The ANP- and BNP-luciferase reporters were described in Grépin et al. (1994). The GATA-4 and -6 expression vectors were described in Durocher et al. (1997). The mouse ET-1 luciferase reporter plasmid was a generous gift of Hiroki Kurihara (University of Tokyo, Japan) and was described in Harats et al. (1995). The GATA-4 and -6 expression vectors were described in Charron et al. (1999) and GATA-2 in Lee et al. (1991). The GATA-5 expression vector was prepared by subcloning the entire coding region into the pCDNA3 vector (*Xho*I/*Kpn*I). The GATA-5 mutant constructs were obtained by PCR-mediated mutagenesis and subcloned in the same manner in pCDNA3. The C-terminal deletion mutant (ΔC) consists of aa 1–300, the N-terminal (ΔN) of aa 180–405, the single carboxy zinc finger mutant of aa 256–405, and the double mutant deletion (ΔCN) retains only the DNA binding domain of aa 180–300. All constructs were confirmed by DNA sequencing.

Electrophoretic mobility shift-assays (EMSA). CV1 cells devoid of GATA-binding activity were transfected with the rat GATA-5 expression vector. 500,000 cells were plated 24 h before transfection with 50 μg of GATA-2, GATA-4, GATA-5, or GATA-5 mutant expression vectors. Nuclear extracts were prepared as previously described (Grépin et al. 1994). Double-stranded oligonucleotides corresponding either to the proximal GATA site of the ANF promoter (Durocher et al. 1997) or the ET-1 GATA element (Harats et al. 1995) were used. Five μg of extracts was incubated with the radiolabeled probe for 15 min at room temperature in a 20-ml volume of 4 mM Tris-HCl(pH = 8), 24 mM EDTA, 0.4 mM MgCl, and 1.6% Ficoll and, where necessary, 100-fold excess of cold unlabeled probes. The samples were loaded on a 4% electrophoresis polyacrylamide gel and run for 3 h at 200 volts.

Western blots. Nuclear extracts (20 μg) of CV1 cells transfected with the different GATA expression vectors were boiled in Laemmli buffer and resolved in SDS-PAGE. Proteins were transferred on Hybond-PVDF membranes and immunoblotted with the Renaissance Chemiluminescence system (NEN Life Sciences, Boston). Anti-HA antibody was purchased from Santa Cruz Biotechnology Inc. (Santa Cruz, Calif.) and used at a dilution of 1/1000.

Genetic mapping of the mouse Gata-5 gene. To establish the chromosomal location of murine *Gata5*, we performed a linkage analysis with The Jackson Laboratory Backcross DNA BSB panel (C57BL/6J × *Mus spretus*) F_1 × C57BL/6J. Genomic DNA of 94 backcross progeny was digested with *Xba*I, electrophoresed in 1% agarose gel, and blotted onto a nylon membrane under standard protocols. Blots were hybridized with a 409-bp ^{32}P-dATP random primer-labeled probe corresponding to nucleotides 807–1215 of the rat Gata-5 cDNA. Hybridization was carried out at 42°C in 50% formamide, 10% dextran sulfate, 1% sodium dodecyl sulfate (SDS), 5× SSC (0.75 M NaCl, 75 mM Na citrate), 2 mM Tris (pH 7.5), 1× Denhardt's solution (0.1% BSA, 0.1% Ficoll, 0.1% polyvinyl pyrrolidone), and 200 μg/ml denatured salmon sperm DNA. Washing was carried out to a final stringency of 0.5 × SSC/1.0% SDS at 60°C. Autoradiography was done at −70°C using Kodak XAR film with the intensification screens; films were developed after exposure for 24–48 h. Detailed mapping data are available at Backcross Panel Mapping Resonace page (http://www.jax.org/resources/documents/cmdata/BSB.htlm).

Results

The GATA-5 protein is conserved between species. Five GATA proteins have been identified so far in mammals: GATA-1, GATA-2, and GATA-3, which are highly restricted to the hematopoietic system, and GATA-4 and -6, which are confined to the heart, the digestive system, and the gonads. The cloning in Xenopus (Jiang and Evans 1996) and in chicken (Laverriere et al. 1994) of GATA-5, the sixth member of the family, whose expression, much like GATA-4 and GATA-6, is restricted to heart and gut, suggested the existence of a subclass of GATA factors that might play important roles in heart and gut development. In order to analyze the evolutionary conservation of GATA-5 and study its role in mammalian heart, we isolated the rat GATA-5 cDNA. Degenerate oligonucleotides corresponding to the highly conserved region of the second zinc finger region of the Xenopus and chicken GATA-5 were synthesized, and polymerase chain reaction (PCR) was carried on cDNA transcribed from RNA of hearts of 3-day-old Sprague Dawley rats. A fragment of 177 bp was cloned in Bluescript vector (pKS): the deduced sequence showed high homology with the corresponding region of the Xenopus and chicken GATA-5 (>98%), compared with 94% and 93% for the same region of rat GATA-4 and rat GATA-6. Two oligonucleotides synthesized according to the sequence read in antisense and sense orientation were used for PCR with two degenerate oligonucleotides corresponding respectively to the start codon (sense orientation) and stop codon (antisense orientation) regions. Two fragments of 806 and 409 bp were cloned in pKS and sequenced. An additional PCR was performed on the two fragments to generate a 1215-bp cDNA corresponding to the entire coding region of GATA-5 encoding a protein of 405 amino acids with a predicted molecular weight of 45 kDa. Computer-assisted sequence analyses showed that the rat GATA-5 protein is highly homologous to its chicken (73%), Xenopus (71%), and the recently identified mouse (97%) counterparts, as illustrated in Figure 1A; the overall homology to the rat GATA-4 and -6 proteins was 60–62%, while the

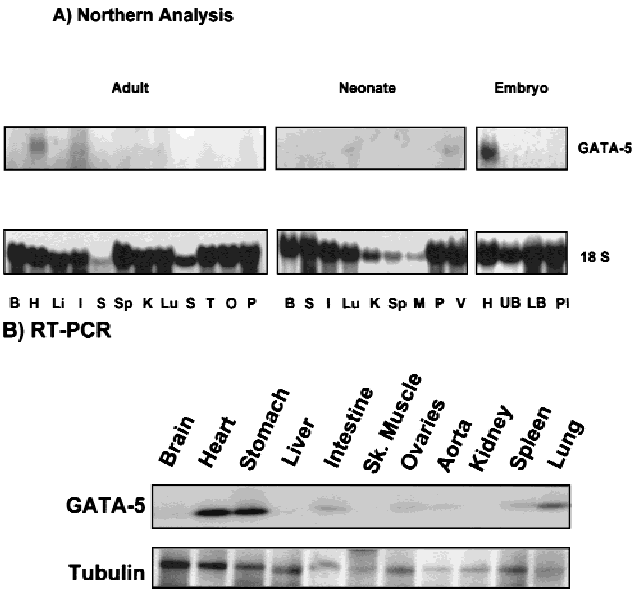

Fig. 2. Tissue distribution of the rGATA-5 mRNA. **(A)** Northern blot analysis: each lane contains 20 μg of total RNA run on formaldehyde agarose gels, transferred to a nylon membrane, and hybridized with a ^{32}P-radiolabelled GATA-5 probe as described in Materials and methods. The radiolabeled 18S probe was used as an internal control. **(B)** RT-PCR amplification of transcripts from various rat adult tissues with specific oligonucleotides as described in Materials and methods, with the amplification of tubulin as an internal control. Products were separated on agarose gel electrophoresis, transferred to a nylon membrane, and hybridized to a specific radiolabeled probe as described in Materials and methods. B (brain), H (heart), I (intestine), L (liver), S (stomach), Lu (lungs), M (skeletal muscle), O (ovaries), T (testes), S (spleen), K (kidney), V (ventricles), UB (upper body), LB (lower body), PL (placenta). Neonates correspond to 3-day-old Sprague Dawley rat. Embryonic tissues were from 14-day mouse embryos.

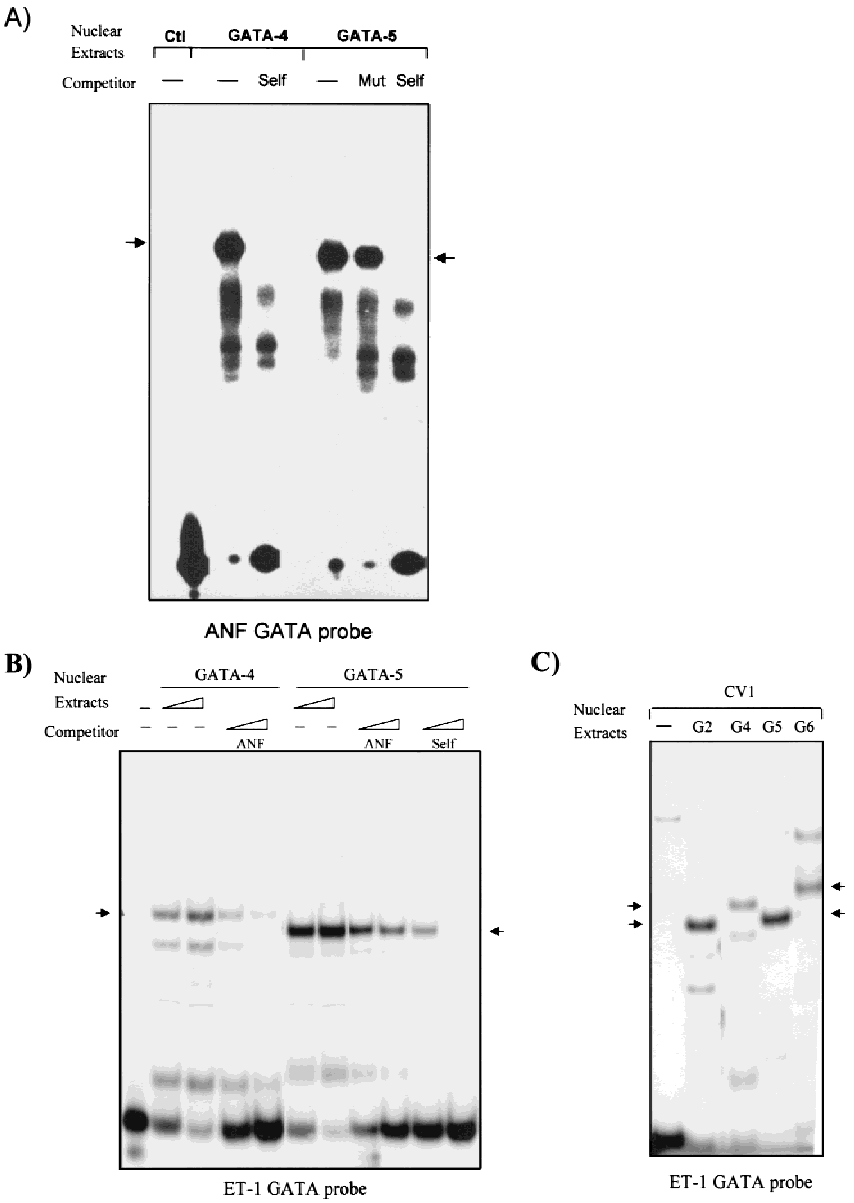

Fig. 3. DNA-binding properties of rGATA-5. Gel shift analyses were carried out with ^{32}P-radiolabeled oligonucleotides corresponding either to the proximal ANF or ET-1 GATA elements. **(A)** 3 μg of nuclear extracts from CV1 cells untransfected (Ctl) or transfected with GATA-4 or GATA-5 expression vectors were used with or without (−) 100-fold excess of unlabeled probe (Self), or a mutant form that no longer binds GATA factors (Mut). **(B)** 3 or 5 μg of nuclear extracts from CV1 cells transfected with GATA-4 or GATA-5 expression vectors were incubated with the ^{32}P-labeled ET-1 GATA element. Competitions were carried out with 25 × and 100 × molar excess of either unlabeled ANF GATA probe (ANF) or the ET-1 GATA element (Self). **(C)** 5 μg of nuclear extracts from CV1 cells transfected with the GATA-2, -4, -5, and -6 expression vectors were incubated with 40,000 cpm of ^{32}P-radiolabeled ET-1 GATA oligonucleotide. Arrows indicate specific GATA binding complexes.

homology in the zinc finger domain was much higher, around 94% (Fig. 1B).

GATA-5 transcripts are expressed at highest levels in the developing heart. Expression of the rat *Gata5* gene was analyzed by Northern blots and RT-PCR. A radiolabeled ^{32}P GATA-5 probe was used on 20 μg of total RNA extracted from different tissues of neonate and adult Sprague Dawley rats, as well as embryos at day 14 of gestation. As shown in Fig. 2A, a unique band of ~2.0 kb was detected in embryonic hearts and in cardiac and extracardiac postnatal tissues. As previously reported for xenopus and chicken (Kelly et al. 1995; Laverriere et al. 1994), expression of GATA-5 in the heart decreased sharply between embryonic and postnatal, suggesting that GATA-5 is mainly an embryonic cardiac regulator (note that the time of exposure for the adult blot is 7 days, whereas that of the embryonic blot is 2 days at −70°C).

In order to better detect low abundant GATA-5 transcripts in postnatal tissues, RT-PCR analysis was carried out. A 409-bp fragment corresponding to nucleotides 806–1215 was amplified with two specific oligonucleotide primers, and the PCR products were further confirmed by Southern blots with a radiolabeled GATA-5 DNA probe. A 546-bp rat tubulin fragment was amplified as internal control. The results showed that, besides the heart, GATA-5 transcripts are found in stomach, intestine, and lung (Fig. 2D). This is similar to the reported tissue distribution of GATA-5 transcripts in adult mouse (Morrisey et al. 1997a) and in postnatal avian (Laverriere et al. 1994) tissues.

Functional properties of the rat GATA-5 protein. In order to perform functional analysis of the GATA-5 protein, the cDNA fragment containing the entire coding region was cloned in the eukaryotic expression vector (pCDNA3) downstream of the CMV promoter and transfected into CV1 cells that lack GATA-binding activity. Nuclear extracts from CV1 cells transfected with this vector or with the corresponding GATA-4 vector were tested for binding to the well-characterized GATA element of the ANF proximal promoter (Charron et al. 1999). As shown in Fig. 3A, both GATA proteins formed a specific complex over this site. Because Gata5 expression is predominant in endocardial cells, we also analyzed the interaction of GATA-5 with the GATA element of the ET-1 promoter. ET-1 is a peptide hormone with potent vasoconstrictive activity synthesized in endothelial and endocar-

dial cells; the ET-1 promoter contains a GATA element and was shown to be transactivated by the endothelial GATA-2 factor (Lee et al. 1991). When used in EMSAs, the ET-1 GATA displayed preferential binding to GATA-5 compared with GATA-4 (Fig. 3B), while the ANF GATA element bound equally well both GATA proteins (Fig. 3A), in accordance with previous studies (Charron et al. 1999). In fact, the ET-1 GATA element had at least three times higher affinity for GATA-5 than the ANF GATA site, as revealed from competition assays (Fig. 3B). Interestingly, the affinity of the ET-1 GATA element was highest for GATA-2 and -5 (Fig. 3C). The transcriptional properties of GATA-5 were also assessed by contransfection with reporter constructs under the control of the ANF, BNP, or ET-1 promoters. The results show that GATA-5 was as effective as GATA-4 in activating the three GATA-dependent promoters (Fig. 4). The highest GATA-5 effect was observed on the ET-1 promoter (20×), which appeared to be a preferential GATA-5 and GATA-2 target, a result that parallels the differential affinities of the various GATA proteins to the ET-1 GATA site. This raises the possibility that GATA-5 may have DNA-binding and functional properties that are distinct from the other GATA factors.

In order to map the functional domains in the GATA-5 protein, we generated four mutant constructs that delete either the N- or C-terminal domain or both, or a truncated form with only the carboxy zinc finger and the carboxy terminal region; none of these mutants alter the nuclear localization signal (NLS). All GATA-5 proteins were hemagglutinin (HA) tagged to allow for immunodetection. As shown in Fig. 5B, all GATA-5 proteins were properly produced from the expression vectors, as assayed by Western blots with an anti-HA antibody. Both wild-type and mutated proteins interacted specifically with a GATA probe (Fig. 5C). However, in cotransfection assays, the C-terminal domain appeared critical for transactivation, as its deletion resulted in a four-fold decrease in ET-1 promoter activation, both in CV1 and in HeLa cells (Fig. 5D). In contrast, the N-terminal deletion mutant was as efficient as the full-length protein in transactivation. Interestingly, the truncated GATA-5 protein that retained the C-terminal region and the carboxy zinc finger was able to bind DNA and activate transcription, suggesting that the N-terminal region and the first zinc finger are dispensable in some cases (the lower degree of activation may be due to differences in expression levels, as suggested by Fig. 5B). The pattern of activity of the various GATA-5 mutants was similar when tested on the ANF and BNP promoters (Fig. 5E). These results map the activation domain for GATA-5 to the C-terminal region, which may be unique to GATA-5 since the activation domains for GATA-1, -2, and -3 were found in the N-terminal region (Visvader et al. 1995), and in the case of GATA-4, two C- and N-terminal domains were mapped (Durocher et al. 1997).

Chromosomal localization of the mouse gata-5 gene. Using a ^{32}P-radiolabeled C-terminal (409-bp) rat cDNA probe, an informative RFLP was observed after *Xba*I digestion of genomic DNA obtained from inbred mouse strains used to create the mapping panel. Parental alleles were identified as specific bands at 9.5 kb (C57BL/6J) and 6.5 kb (*Mus spretus*). The location of mouse *Gata5* was determined by comparing the inheritance of the *Mus spretus* allele among 93 backcross progeny with 890 loci previously mapped in this cross. Gene order was established by minimizing the number of double recombination events. The results indicate that *Gata5* maps to distal Chr 2, cosegregating with three microsatellites, *D2Mit74, D2Ucl18,* and *D2Ucl29* (Fig. 6). Two recombinants were detected between *Gata5* and a cluster of markers including another microsatellite, *D2Ucl1,* as well as the Gs protein alpha stimulatory regulator gene (*Gnas*), the phosphoenolpyruvate carboxykinase 1 gene (*Pck1*), and the synaptobrevin-like 1 pseudogene (*Sybl1-ps*). Two mouse mutants are known to map to the

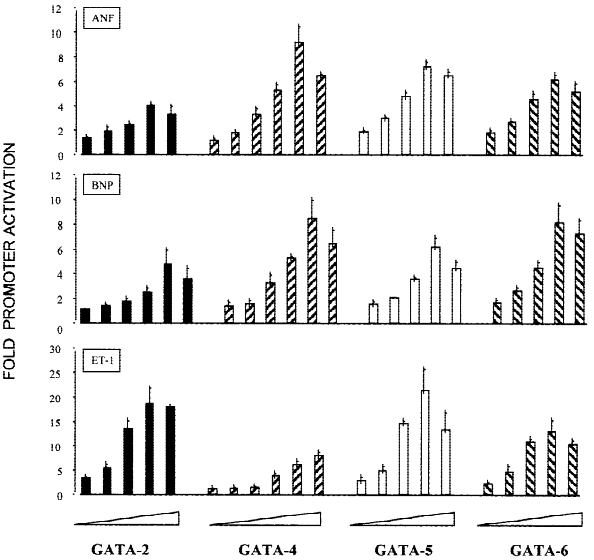

Fig. 4. Transactivation properties of the recombinant rGATA-5 protein. In all cotransfections, the amounts of the reporter genes (ANF, BNP, or ET-1-luc) were kept constant at 3 μg. The data shown are the mean of two to four independent experiments, each carried out in duplicate in CV1 cells. The results are expressed as fold activation relative to cotransfection with the backbone vector; 25–500 ng of GATA-2, -4, -5, and -6 expression vectors were used. Similar results were also obtained in HeLa cells. Note the different scales used and the higher sensitivity of the ET-1 promoter for GATA-5 and -2 relative to GATA-4.

telomeric region of Chr 2 in the vicinity of *Gata5*, ragged (*ra*) and wasted (*wst*) (1). With the exception of edema in *ra*/*ra* homozygotes, however, defects involving the cardiac or gastrointestinal system that would be consistent with a *Gata5* mutation have not been described in association with either mutation. The human orthologue of mouse *Gata5* is predicted to map to Chr 20q13 (24) and the rat *Gata5* gene to Chr 3q43 (http://www.informatics.jax.org/homology.html)

Discussion

In this paper, we report the cloning and characterization of rat GATA-5, a member of the GATA family of transcription factors expressed in the developing heart. While GATA-4, -5, and -6 transcripts are present in the heart, GATA-5 mRNAs are the least abundant (5- to 10-fold less than GATA-4), suggesting that only specific subsets of cells express GATA-5 in the heart. Indeed, in situ hybridization studies show that, both in Xenopus and rodent, GATA-5 mRNA is found predominantly in the endocardium as well as the outflow-tract (Kelley et al. 1993; Morrisey et al. 1997a; and our unpublished data). Except for $NFAT_c$ (de la Pompa et al. 1998), few transcription factors that are implicated in the differentiation of the endocardial endothelial cells are known. Endocardial differenation is accompanied by the expression of specific genes such as the homeobox gene *msx-1* (Chan-Thomase et al. 1993; Eisenberg and Markwald 1995; Shen et al. 1994) fibulin, fibronectin and tenascin-X (Bristow 1995; Eisenberg and Markwald 1995). The presence of conserved GATA motifs in some of these genes, like *msx-1* (Shen et al. 1994), tenascin-x (Bristow et al. 1993), and endothelin-1 (Wilson et al. 1990), suggests a role for GATA-5 in regulating endocardial genes and thus in cardiac lineage differentiation. Interestingly, some myocardial cells coexpress GATA-4 and GATA-5 in the embryonic heart (G. Nemer and M. Nemer, unpublished data). Given that GATA-5 is as potent an activator of cardiac promoters as GATA-4, it may well be able to substitute—at least in part—for GATA-4 with respect to cardiac transcription; this may in turn be relevant for understanding the

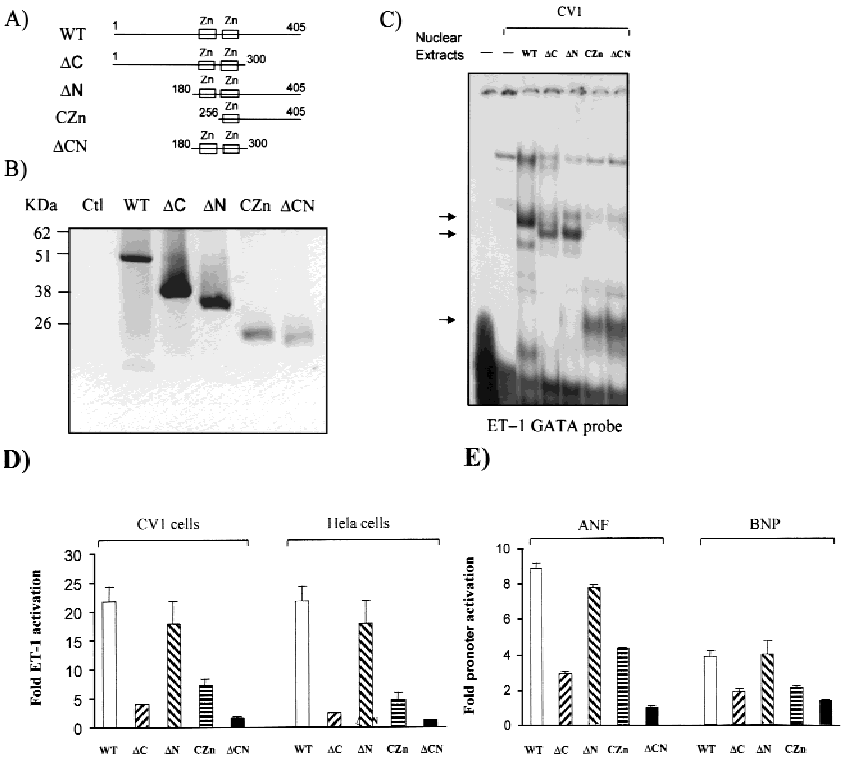

Fig. 5. Structure-function analysis of the rGATA-5 protein. **(A)** Schematic representation of the different GATA-5 constructs. **(B)** Western blot analysis of the different GATA-5 recombinant proteins. Nuclear extracts from CV1 cells untransfected (−) or transfected with the different GATA-5 tagged constructs (WT, ΔN, ΔC, CZn, and ΔCN) were used. Revelation of the tagged proteins run on an SDS-PAGE was carried out with an anti-HA antibody. **(C)** Gel shift analysis confirms that all the different recombinant proteins are able to bind GATA elements (specific complexes shown by arrows). **(D)** 100 ng of each of the full-length (WT) GATA-5 construct, the N- and C-terminal mutant deletion, or both (ΔN, ΔC, ΔCN) or the single carboxy zinc finger and the carboxy domain (CZn) were used in all experiments in CV1 and Hela cells with 3 μg of the ET-1-luciferase reporter gene. The results are expressed as fold activation relative to cotransfection with the backbone vector. **(E)** Cotransfection assays in CV1 cells were carried out with 3 μg of the ANF and BNP reporter genes and 100 ng of GATA-5 expression. The results are the mean of three independent experiments, each in duplicate.

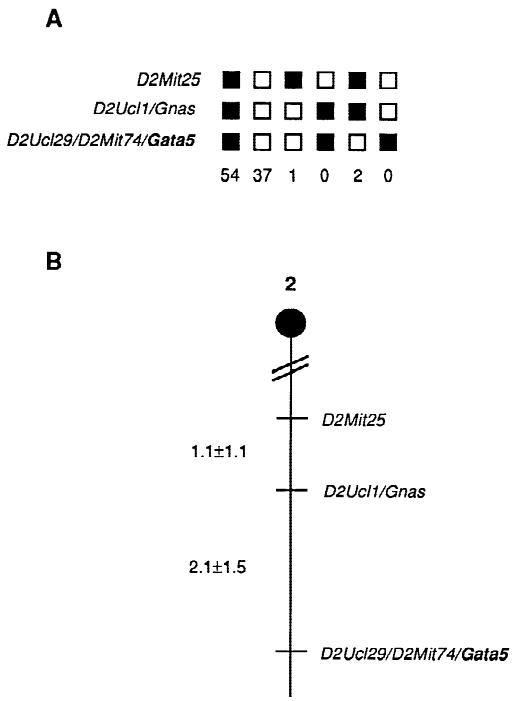

Fig. 6. Chromosomal mapping of the mouse *Gata5* gene. **(A)** Haplotypes observed in 94 backcross mice segregating *Gata5* with Chr 2 markers. Open boxes represent homozygous progeny; closed boxes represent heterozygous progeny. The number of animals exhibiting each haplotype is shown at the bottom of each column. **(B)** A partial linkage map of mouse Chr 2 showing the location of *Gata5* in relation to flanking markers. Recombination distances in centimorgans (±SE) are shown to the left of the chromosomal map.

presence of many markers of cardiomyocyte differentiation in the cardiac crescent of GATA-4 null embryos (Molkentin et al. 1997). The results presented also revealed that GATA-5 may have preferential downstream targets in cells coexpressing other GATA factors. Indeed, the ET-1 promoter was more sensitive to GATA-5 than -4 and was induced three times more by similar levels of GATA-5 and -4 (21 × vs 7 ×), possibly a result of differential affinity of the two proteins for the ET-1 GATA elements. In contrast, the ANF and BNP promoters were activated similarly by the three cardiac GATA factors. Thus, GATA-4, -5, and -6 could substitute for each other in activation of some but not all cardiac genes; this in turn suggests that they could only partially compensate for each other in heart formation.

Interestingly, the major activation domain of GATA-5 mapped to the C-terminal region. In contrast, the major activation domains for GATA-1, -2, and -3 were found in N-terminal domain (Visvader et al. 1995), whereas two independent activation domains were mapped in the C- and N-terminal domains of GATA-4 (Morrisey et al. 1997b). In this respect it is noteworthy that the C-terminal region of GATA-5 proteins is highly conserved across species, suggesting evolutionary pressure to preserve an important function. For example, a 60-amino acid C-terminal domain located between aa 366 and 405 is 70% identical among Xenopus, chicken, and rodent GATA-5. Whether this domain interacts with other cofactors, such as members of the NK2 family, will be interesting to investigate. Indeed, GATA-5 was shown to cooperate with the cardiac-specific Nkx2-5 protein, probably through the C-terminal domain (Durocher et al. 1997). Other members of the NK2 family are also expressed in similar tissues as GATA-5; this includes Nkx2-3 and Nkx2-6 which are found in the gut (Biben et al. 1998; Buchberger et al. 1996), and Nkx2-1, which is expressed among others in lung (Lazzaro et al. 1991). Thus, it is tempting to speculate that GATA-5 might play important roles in these tissues through cooperative interaction with other tissue-specific factors.

Finally, the mouse *Gata-5* gene maps to Chr 2, whereas the

Gata4 and *Gata6* genes are localized on mouse Chrs 14 and 18, respectively (Qureshi et al. 1996; Sebastiani et al. 1995). Thus, despite extensive sequence homology, these genes appear to have been duplicated from an ancestral precursor early in evolution.

Acknowledgments. We are grateful to Lise Laroche for secretarial assistance, to Daniel Durocher for help with the isolation of a GATA-5 cDNA fragment, and to other members of the Nemer laboratory for reagents and helpful discussions. This work was supported by research grants from the Medical Research Council of Canada (MRC). G. Nemer holds a Gibco-IRCM studentship. S. Qureshi is a recipient of a Clinician Scientist Award from MRC and a Career Award in the Biomedical Sciences from the Burroughs Welcome Fund. D. Malo is a scholar of Fonds de la recherche en santé du Québec (FRSQ), and M. Nemer is an MRC scientist.

References

Biben C, Hatzistavrou T, Harvey RP (1998) Expression of NK-2 class homeobox gene Nkx2-6 in foregut endoderm and heart. Mech Dev 73, 125–127

Bristow J (1995) The search for genetic mechanisms of congenital heart disease. Cell Mol Biol Res 41, 307–319

Bristow J, Tee MK, Gitelman SE, Mellon SH, Miller WL (1993) Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene overlapping P450c21B. J Cell Biol 122, 265–278

Buchberger A, Pabst O, Brand T, Seidl K, Arnold HH (1996) Chick NKx-2.3 represents a novel family member of vertebrate homologous to the Drosophila homeobox gene tinman—differential expression of cNKx-2.3 and cNKx-2.5 during heart and gut development. Mech Dev 56, 151–163

Chan-Thomas PS, Thompson RP, Robert B, Yacoub MH, Barton PJ (1993) Expression of homeobox genes Msx-1 (Hox-7) and Msx-2 (Hox-8) during cardiac development in the chick. Dev Dyn 197, 203–216

Charron F, Paradis P, Bronchain O, Nemer G, Nemer M (1999) Cooperative interaction between GATA-4 and GATA-6 regulates myocardial ANF gene expression. Mol Biol Cell, 19, 4355–4365

de la Pompa JL, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, et al. (1998) Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. Nature 392, 182–186

Durocher D, Charron F, Warren R, Schwartz RJ, Nemer M (1997) The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. EMBO J 16, 5687–5696

Eisenberg LM, Markwald RR (1995) Molecular regulation of atrioventricular valvuloseptal morphogenesis. Circ Res 77, 1–6

Grépin C, Dagnino L, Robitaille L, Haberstroh L, Antakly T et al. (1994) A hormone-encoding gene identifies a pathway for cardiac but not skeletal muscle gene transcription. Mol Cell Biol 14, 3115–3129

Grépin C, Robitaille L, Antakly T, Nemer M (1995) Inhibition of transcription factor GATA-4 expression blocks in vitro cardiac muscle differentiation. Mol Cell Biol 15, 4095–4102

Grépin C, Nemer G, Nemer M (1997) Enhanced cardiogenesis in embryonic stem cells overexpressing the GATA-4 transcription factor. Development 124, 2387–2395

Harats D, Kurihara H, Belloni P, Oakley H, Ziober A, et al. (1995) Targeting gene expression to the vascular wall in transgenic mice using the murine preproendothelin-1 promoter. J Clin Invest 95, 1335–1344

Heikinheimo M, Scandrett JM, Wilson DB (1994) Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. Dev Biol 164, 361–373

Ip HS, Wilson DB, Heikinheimo M, Tang Z, Ting CN et al. (1994) The GATA-4 transcription factor transactivates the cardiac muscle-specific troponin C promoter-enhancer in nonmuscle cells. Mol Cell Biol 14, 7517–7526

Jiang YM, Evans T (1996) The Xenopus GATA-4/5/6 genes are associated with cardiac specification and can regulate cardiac-specific transcription during embryogenesis. Dev Biol 174, 285–270

Kelley C, Blumberg H, Zon LI, Evans T (1993) GATA-4 is a novel transcription factor expressed in endocardium of the developing heart. Development 118, 817–827

Kelly R, Alonso S, Tajbakhsh S, Cossu G, Buckingham M (1995) Myosin light chain 3F regulatory sequences confer regionalized cardiac and skeletal muscle expression in transgenic mice. J Cell Biol 129, 383–396

Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JB et al. (1994) GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. J Biol Chem 269, 23177–23184

Lazzaro D, Price M, De Felice M, Di Lauro R (1991) The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. Development 113, 1093–1104

Lee ME, Temizer DH, Clifford JA, Quertermous T (1991) Cloning of the GATA-binding protein that regulates endothelin-1 gene expression in endothelial cells. J Biol Chem 266, 16188–16192

Molkentin JD, Kalvakolanu DV, Markham BE (1994) Transcription factor GATA-4 regulates cardiac muscle-specific expression of the a-myosin heavy-chain gene. Mol Cell Biol 14, 4947–4957

Molkentin JD, Lin Q, Duncan SA, Olson EN (1997) Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Genes Dev 11, 1061–1072

Morrisey EE, Ip HS, Tang ZH, Lu MM, Parmacek MS (1997a) GATA-5—a transcriptional activator expressed in a novel temporally and spatially restricted pattern during embryonic development. Dev Biol 183, 21–36

Morrisey EE, Ip HS, Tang Z, Parmacek MS (1997b) GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily. J Biol Chem 272, 8515–8524

Pandolfi PP, Roth ME, Karis A, Leonard MW, Dzierzak E et al. (1995) Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. Nat Genet 11, 40–44

Pevny L, Lin CS, D'Agati V, Simon MC, Orkin SH (1995) Development of hematopoietic cells lacking transcription factor GATA-1. Development 121, 163–172

Qureshi ST, Bronchain O, Nemer M, Malo D (1996) Mapping of the *Gata6* gene to mouse chromosome 18. Mamm Genome 7, 705–706

Sebastiani G, Grépin C, Nemer M, Malo D (1995) The mouse *Gata4* transcription factor maps to chromosome 14. Mamm Genome 6, 442–445

Shen R, Chen Y, Huang L, Vitale E, Solursh M (1994) Characterization of the human MSX-1 promoter and an enhancer responsible for retinoic acid induction. Cell Mol Biol Res 40, 297–312

Tamura S, Wang XH, Maeda M, Futai M (1993) Gastric DNA-binding proteins recognize upstream sequence motifs of parietal cell-specific genes [published erratum appears in Proc Natl Acad Sci USA 1994 May 10;91(10):4609]. Proc Natl Acad Sci USA 90, 10876–10880

Tsai FY, Keller G, Kuo FC, Weiss M, Chen J et al. (1994) An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature 371, 221–226

Visvader JE, Crossley M, Hill J, Orkin SH, Adams JM (1995) The C-terminal zinc finger of GATA-1 or GATA-2 is sufficient to induce megakaryocytic differentiation of an early myeloid cell line. Mol Cell Biol 15, 634–641

Weiss MJ, Keller G, Orkin SH (1994) Novel insights into erythoid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. Genes Dev 8, 1184–1197

Wilson DB, Dorfman DM, Orkin SH (1990) A non-erythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. Mol Cell Biol 10, 4854–4862