Direct transcriptional regulation of Gata4 during early endoderm specification is controlled by FoxA2 binding to an intronic enhancer

Anabel Rojas1,*, William Schachterle2, Shan-Mei Xu2, Franz Martín1, and Brian L. Black2,3

1 Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER) -CIBERDEM, Avda. Americo Vespucio s/n, 41092, Sevilla, SPAIN 2 Cardiovascular Research Institute and 3 Department of Biochemistry and Biophysics University of California, San Francisco, CA 94143-2240, USA

Abstract

The embryonic endoderm is a multipotent progenitor cell population that gives rise to the epithelia of the digestive and respiratory tracts, the liver and the pancreas. Among the transcription factors that have been shown to be important for endoderm development and gut morphogenesis is GATA4. Despite the important role of GATA4 in endoderm development, its transcriptional regulation is not well understood. In this study, we identified an intronic enhancer from the mouse Gata4 gene that directs expression to the definitive endoderm in the early embryo. The activity of this enhancer is initially broad in all endodermal progenitors, as demonstrated by fate mapping analysis using the Cre/loxP system, but becomes restricted to the dorsal foregut and midgut, and associated organs such as dorsal pancreas and stomach. The function of the intronic Gata4 enhancer is dependent upon a conserved Forkhead transcription factor-binding site, which is bound by recombinant FoxA2 in vitro. These studies identify Gata4 as a direct transcriptional target of FoxA2 in the hierarchy of the transcriptional regulatory network that controls the development of the definitive endoderm.

Keywords

GATA4; Forkhead proteins; FoxA2; endoderm progenitors; enhancer; transgenic mouse

Introduction

During gastrulation, the developing embryo undergoes a precise morphogenetic program that leads to the establishment of the basic body plan. Just before gastrulation, the amniote embryo consists of an epithelial layer called the epiblast, which is covered by an extraembryonic layer, known as the visceral endoderm (Lawson et al., 1986; Tam and Loebel, 2007). During gastrulation, cells in the epiblast undergo an epithelial-to-mesenchymal transition, and ingress through the primitive streak (Lawson et al., 1991).
Ingressing epiblast cells form the mesoderm and endoderm while the embryonic surface forms the ectoderm. The presumptive definitive endoderm, which gives rise to endodermal organs, displaces the extraembryonic visceral endoderm, which gives rise to extraembryonic structures such as the yolk sac (Kadokawa et al., 1987; Weber et al., 1999). The visceral endoderm shares the expression of many genes with the definitive endoderm (Sherwood et al., 2007).

After gastrulation, a series of morphogenetic movements transform the definitive endoderm into a primitive gut tube surrounded by mesoderm. The anterior and posterior regions of the primitive tube fold ventrally and the tube becomes regionalized along the dorsal-ventral and anterior-posterior axes into foregut, midgut, and hindgut. These subregions are defined by specific genetic programs controlled by cell non-autonomous factors derived from nearby mesoderm and ectodermal tissues and mediated by transcription factors that directly regulate gene expression in different endoderm-derived organs. (Bort et al., 2004; Grapin-Botton and Melton, 2000; Kumar et al., 2003; Molotkov et al., 2005; Serls et al., 2005; Stainier, 2002; Tremblay and Zaret, 2005; Wells and Melton, 2000). The ventral foregut gives rise to the liver, ventral pancreas, and lungs, while dorsal foregut and anterior regions of the midgut contribute to the dorsal pancreas, stomach, esophagus, and duodenum. The midgut and hindgut form the small and large intestines, respectively (Lawson et al., 1986).

Among the transcription factors that have been shown to be crucial for endoderm development is FoxA2, a member of the Forkhead family of transcription factors (Costa et al., 1989; Lai et al., 1990). In the early embryo, FoxA2 is expressed in both visceral and definitive endoderm, the node, the notochord, and the floorplate (Ang et al., 1993; Monaghan et al., 1993; Ruiz i Altaba et al., 1993; Sasaki and Hogan, 1993). Later in development, FoxA2 is also present in endoderm-derived tissues, including liver, lung, stomach, pancreas, small intestine, and colon (Besnard et al., 2004). Targeted disruption of foxa2 leads to embryonic lethality, and foxa2-null embryos display malformation of the node, notochord, and foregut endoderm (Ang et al., 1993; Weinstein et al., 1994). Recent studies, using conditional inactivation approaches, have revealed that FoxA2 is also required at later stages of development in various endoderm-derived organs (Gao et al., 2008; Kaestner, 2005; Lantz et al., 2004; Lee et al., 2005; Wan et al., 2004; Zhang et al., 2005). Mechanistically, FoxA2 functions as a “pioneer factor” by displacing linker histones from compacted chromatin and facilitating the binding of other transcription factors, including the zinc finger transcription factor GATA4 (Gualdi et al., 1996; Zaret, 1999; Zaret, 2008; Zaret et al., 2008).

GATA4 belongs to a family of transcription factors involved in the differentiation of the endoderm in several evolutionarily diverse organisms (Rehorn et al., 1996; Reiter et al., 2001; Zhu et al., 1997). Among the members of the GATA family, GATA4 is the first to be expressed in the visceral and definitive endoderm. GATA4 is also expressed broadly in the mesoderm of the early mouse embryo (Arceci et al., 1993; Molkentin, 2000). Homozygous inactivation of Gata4 in mice leads to embryonic lethality, and Gata4-null embryos display severe defects in the visceral endoderm, malformations in the ventral foregut, and major disruptions in heart morphogenesis (Kuo et al., 1997; Molkentin et al., 1997), establishing a crucial role of this transcription factor in these cell lineages. More recently, conditional inactivation, tetraploid complementation, and mosaic analyses of Gata4 mutant mouse embryos have also revealed its important role in the differentiation of a variety of endodermal-derived tissues (Battle et al., 2008; Bosse et al., 2006; Dusing et al., 2001; Jacobsen et al., 2002; Watt et al., 2007). Similarly, morpholino knockdown experiments demonstrated a requirement of GATA4 for the proper formation of endodermal organs in the zebrafish (Holtzinger and Evans, 2005). In addition to its role in differentiation of endodermal derivatives, GATA4 also appears to participate in endoderm specification in...
cooperation with FoxA2 (Bossard and Zaret, 1998; Cirillo et al., 2002; Zaret, 1999). However, despite the importance of GATA4 in the development of the early endoderm, relatively little is known about its transcriptional regulation in vivo.

In this study, we identified a transcriptional enhancer from the mouse Gata4 gene, referred to as Gata4 G4, which is located in the second intron of the mouse Gata4 gene. Importantly, this novel intronic enhancer directs expression to the definitive endoderm from early stages in mouse development. As development proceeds, the activity of the Gata4 G4 enhancer becomes restricted to cells of the dorsal foregut and midgut. The cells marked by the activity of this novel Gata4 enhancer at early stages contribute to all endodermal progenitors, as indicated by Cre-based lineage tracing experiments. Reporter gene expression directed by the Gata4 G4 enhancer in transgenic embryos in vivo matches the expression of transcripts for the Forkhead transcription factor FoxA2, and we show the Gata4 G4 enhancer contains an essential conserved Forkhead binding site that is efficiently bound by recombinant FoxA2 protein and is required for enhancer activity in vivo. Thus, our studies identify Gata4 as a direct transcriptional target of the Forkhead transcription factor FoxA2 in the early endoderm and establish a Gata4-expressing progenitor population that contributes to the entire definitive endoderm.

Materials and Methods

Bioinformatic analyses, cloning, and mutagenesis

Sequence comparisons were performed by using BLAST and VISTA algorithms (Altschul et al., 1990; Mayor et al., 2000). The 1,107-bp G4 fragment of the mouse Gata4 gene was generated by PCR from mouse genomic DNA using the following primers: 5′-tgctgaggaacgaagcttccag-3′ and 5′-ccctgtacttcagctgg-3′. This fragment was then cloned as a HindIII-PstI fragment into the HindIII-PstI sites of the transgenic reporter plasmid Hsp68-lacZ (Kothary et al., 1989). The Forkhead site mutant in Gata4 G4 was created by using primers with the following plus strand sequence: 5′-attataatcacagactagtcgccagaagtctcaaag-3′. The sequence of the mutant fragment was confirmed by sequencing on both strands. The Gata4 G4 fragment fused to the Gata4 minimal promoter containing the mouse Gata4 sequences from −443 to +52, relative to the transcriptional start site, was cloned into a Cre expression plasmid containing the Cre cDNA and the SV40 splice and poly A signal sequences to generate plasmid Gata4-G4Gp-Cre. The GenBank Accession number for the sequence of Gata4 G4 endodermal enhancer is HM212783.

Generation of transgenic mice

Transgenic lacZ reporter and Cre constructs were digested from their parental plasmid backbones with SalI, gel purified, and suspended in 5 mM Tris-HCl, 0.2 mM EDTA, pH 7.4, at a concentration of 2 μg/ml for pronuclear injection as described previously (Hogan et al., 1994). Injected embryos were implanted into pseudopregnant FVB females, and embryos were collected at indicated times for F0 analysis or were allowed to develop to adulthood for the establishment of stable transgenic lines. DNA was extracted from the yolk sac of embryos or from tail biopsies as previously described (Dodou et al., 2003). The presence of the Gata4-G4-lacZ and Gata4-G4Gp-Cre transgenes were detected by Southern blot using specific probes for each allele.

X-gal staining, immunohistochemistry, and in situ hybridization

β-galactosidase expression in lacZ transgenic embryos or tissues was detected by X-gal staining, which was performed as described previously (Dodou et al., 2003). For the analysis of the Gata4-G4Gp-Cre line, transgenic male founders were crossed with female Rosa26R.
lacZ reporter mice (Soriano, 1999) and embryos were collected at different stages of development. Transverse and sagittal sections from X-gal stained embryos and tissues were prepared and counterstained with Neutral Fast Red as described previously (Anderson et al., 2004). GATA4 immunohistochemistry was performed as previously described (Rojas et al., 2009). Briefly, dewaxed sections were boiled in antigen retrieval and incubated with 3% H$_2$O$_2$ for 15 min prior to block with 3% normal goat serum. Incubation with anti-GATA4 antibody (1:300 dilution) was done overnight at 4°C. Incubation with biotinylated anti-goat antibody (Vector Laboratories) at a 1:300 dilution was done at room temperature for 1 h. Immunoperoxidase staining was performed using the Vectastain Elite ABC kit (Vector Laboratories) and developed using the peroxidase substrate DAB (Vector Laboratories).

Whole mount in situ hybridization was performed as described previously (Rojas et al., 2005). Following staining, embryos were sectioned at a thickness of 7 μm and counterstained with Neutral Fast Red for visualization of embryonic structures. foxa2 antisense probe was generated from pBluescript-FoxA2 containing the full-length cDNA linearized with BamHI, and transcribed using T3 polymerase.

**Electrophoretic mobility shift assay (EMSA)**

DNA binding reactions were performed as described previously (Dodou et al., 2003). The foxa2 cDNA was transcribed and translated using the TNT Coupled Transcription-Translation System (Promega), as described in the manufacturer’s directions. FoxA2 protein was generated from pcDNA1-FoxA2 plasmid using T7 polymerase. The sense strand sequence of the mouse Gata4 G4 Forkhead site used for EMSA was: Fox 5′-gagggtaatatactacaaaaagcttaagat-3′. The sense strand sequences for the Forkhead mutant site was the same as for the mutagenic primers described above.

**Chromatin Immunoprecipitation (ChIP)**

ChIP assays were performed using the ChIP assay kit from Upstate Pharmaceuticals following the recommendations of the manufacturer. Briefly, a 10 cm plate containing approximately 1×10$^6$ AR42J rat cells (Rosewicz et al., 1992), grown in Dulbecco modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), were treated with 1% paraformaldehyde at 37°C for 10 min to crosslink protein-DNA complexes. Cells were then lysed and sonicated to shear the DNA into fragments approximately 500 bp in size. The cleared supernatant was divided into two samples. One of the samples was incubated with 4 μg of anti-FoxA2 antibody (Santa Cruz Biotechnology) and the other sample was incubated with 4 μg of anti-rabbit IgG (Santa Cruz Biotechnology) as a nonspecific control overnight at 4°C. The DNA fragments were then precipitated after incubating the lysate and antibody mixture with protein A-agarose beads for 1 h. Reactions were incubated in 200 mM NaCl at 65°C for 4 h to reverse the crosslinks, and DNA was recovered by phenol-chloroform extraction. The following primers were used to amplify a fragment of 290 bp encompassing the Gata4 G4 Forkhead site: 5′-cagttgaattgtcccaaatcagatc-3′ and 5′-ccaactgtggatgaacttagcttcc-3′.

**Results**

An intronic Gata4 endoderm-specific enhancer directs expression to the dorsal foregut and midgut in mouse transgenic embryos

Comparison of the Gata4 sequences of mouse, human, and opossum identified nine major regions of conserved non-coding sequences, referred to as G1–G9 (Fig. 1). Previously, we demonstrated that two of these conserved regions function as independent, modular transcriptional enhancers: Gata4 G2 functions as a specific enhancer in lateral mesoderm and a subset of its derivatives in the liver mesenchyme and Gata4 G8 functions as a late-
acting endoderm enhancer (Rojas et al., 2005; Rojas et al., 2009). The \textit{Gata4} G8 enhancer directs expression to the foregut and midgut in transgenic mouse embryos, beginning at embryonic day (E) 9.5 (Rojas et al., 2009), but this enhancer is not active at earlier embryonic stages when endogenous \textit{Gata4} transcripts are already robustly expressed (Molkentin, 2000; Nemer and Nemer, 2003). Together, these observations suggested that additional \textit{Gata4} endoderm enhancers must exist and prompted us to test additional conserved sequences from the \textit{Gata4} locus for enhancer activity in early transgenic embryos.

To identify additional \textit{Gata4} endoderm enhancers, we cloned each conserved noncoding region from the \textit{Gata4} locus individually into the transgenic reporter plasmid \textit{Hsp68-lacZ} (Kothary et al., 1989) and tested for enhancer activity in transgenic mouse embryos. These analyses indeed resulted in the identification of an additional endoderm-specific enhancer element, which we refer to as \textit{Gata4} G4, with early activity in the visceral and definitive endoderm (Fig. 1). The \textit{Gata4} G4 enhancer element encompasses a conserved region of 1,107 bp, located between exons 2 and 3 (Fig. 1). To evaluate \textit{Gata4} G4 activity \textit{in vivo}, we generated seven independent transgenic lines harboring the \textit{Gata4-G4-lacZ} transgene and analyzed them for \(\beta\)-galactosidase reporter expression during embryonic development. Six of the transgenic lines showed a consistent \textit{lacZ} expression pattern in all the embryonic stages analyzed. We first observed enhancer activity at E7.5 (Fig. 2A). At that stage, \textit{Gata4-G4-lacZ} directed expression to the visceral endoderm of the yolk sac, in agreement with the previously described early expression of endogenous \textit{Gata4} in the extra-embryonic endoderm (Kuo et al., 1997; Molkentin et al., 1997; Nemer and Nemer, 2003) (Fig. 2A). Extra-embryonic expression was transient and no longer observed at any developmental stage after E7.5 (Fig. 4A and data not shown). Later, at E8.5, robust \textit{Gata4-G4-lacZ} transgene expression was observed in the primitive gut tube formed by the definitive endoderm (Fig. 2B–D). By E9.5, \(\beta\)-galactosidase activity was observed in the ventral and dorsal part of the gut tube, and became clearly restricted to the midgut (Fig. 2E, F, K). In the foregut, \textit{lacZ} staining was only observed in the most caudal/posterior regions, indicating that enhancer activity did not extend more rostrally than the developing liver (Fig. 2H). The expression of the transgene was completely absent from the mesoderm component of the gut, demonstrating the endodermal specificity of the \textit{Gata4} G4 element \textit{in vivo}.

To characterize the \textit{Gata4} endoderm enhancer in more detail, we compared the expression pattern directed by the enhancer with the expression pattern of the endodermal transcription factor FoxA2 (Ang et al., 1993). The expression patterns of the \textit{Gata4-G4-lacZ} transgene and \textit{foxa2} appeared to be nearly identical in the posterior foregut and midgut (Fig. 2, compare panels E, F, H, K to panels G, I, L). The expression of the \textit{Gata4-G4-lacZ} transgene also partially overlapped the expression of endogenous GATA4 protein in the posterior foregut and midgut at E9.5 (Fig. 2 compare panels H, K to panels J, M). Thus, these results indicate that G4 enhancer faithfully recapitulates \textit{Gata4} endogenous expression in the early endoderm in transgenic embryos.

Endogenous GATA4 protein expression was broader than \textit{Gata4-G4-lacZ} and included other derivatives of the endoderm and mesoderm, such as the entire foregut and gut mesenchyme (Fig. 2J, M), consistent with previously published reports (Arceci et al., 1993). The broader expression of endogenous GATA4 protein relative to \textit{Gata4-G4-lacZ} was expected since endogenous GATA4 expression is controlled by multiple independent enhancers with activity in the gut (Rojas et al., 2005; Rojas et al., 2009), and the \textit{Gata4} G4 enhancer controls only a subset of the endogenous pattern of expression. In addition to its expression in the developing gut endoderm, we also observed \textit{Gata4-G4-lacZ} expression in the floor plate (Fig. 2H, K) even though endogenous GATA4 protein was not detectable in the floor plate by immunohistochemistry (Fig. 2J, M). This discrepancy might reflect the
absence of additional regulatory sequences in the *Gata4* G4 enhancer that normally repress endogenous *Gata4* expression in the floor plate.

**Gata4 G4 enhancer activity is restricted to the epithelium of the hind-stomach and dorsal pancreatic bud**

At E11.5, expression directed by the *Gata4* G4 enhancer was observed only in a subset of endoderm-derived organs, the epithelium of the glandular stomach, and the dorsal pancreatic bud (Fig. 3A–C). By E13.5, transgene activity was restricted to the glandular (hind) stomach, although at lower levels compared to earlier stages of development and it was absent in the anterior stomach (Fig. 3D). These results are in agreement with the previously described expression of endogenous *Gata4* in the developing stomach (Jacobsen et al., 2005). In the dorsal pancreas, only a few cells in the epithelium displayed detectable β-galactosidase activity at E13.5 (Fig. 3E). No transgene expression was observed in the mesenchyme surrounding the stomach and pancreas, indicating the specificity of the G4 enhancer to direct expression in the endodermal component of these organs. At later stages in fetal development, the *Gata4-G4-lacZ* transgene was no longer active since no β-galactosidase activity was observed in either the hind stomach epithelium or the pancreas (Fig. 3F). The decreased expression of the transgene coincides with terminal cell differentiation and maturation, suggesting that the *Gata4* G4 enhancer is active mainly at early stages of endoderm development.

**Gata4-G4-Cre expressing cells and their descendents contribute to all endodermally-derived organs**

The results shown in Figure 2 demonstrated that the *Gata4-G4-lacZ* transgene was expressed very early in the definitive endoderm of the mouse embryo but then became restricted to a subset of endoderm derivatives. However, the broad early activity of the *Gata4* G4 enhancer in the endoderm suggested that the cells initially and transiently marked by *Gata4-G4-lacZ* might contribute extensively to organs derived from the definitive endoderm. Therefore, to investigate the developmental potential of the progenitors marked by the early activity of the *Gata4* G4 enhancer, we generated transgenic mice in which the expression of Cre recombinase was controlled by the *Gata4* G4 enhancer linked to the minimal promoter region of *Gata4*. We crossed these mice, referred to as *Gata4-G4Gp-Cre*, to *Rosa26R lacZ* reporter mice, which express β-galactosidase constitutively and indelibly in a Cre-dependent manner (Soriano, 1999). We compared the Cre-dependent *lacZ* expression in these embryos to the *lacZ* expression directly under the control of the *Gata4* G4 enhancer at different stages of development. As mentioned above, *Gata4* G4 enhancer activity in the yolk sac was transient, rapidly extinguishing after E7.5 (Figs. 2A, 4A). However, in *Gata4-G4Gp-Cre^{TG0}; Rosa26^{lacZ/+}* embryos, β-galactosidase activity was observed in the yolk sac at all stages of embryonic development (Fig. 4B. and data not shown). These results indicate that the early activation of the *Gata4* G4 enhancer in the visceral endoderm marks all of the progenitors that will give rise to the extra-embryonic endoderm of the yolk sac. As shown in Figure 3, at E11.5, the *Gata4-G4-lacZ* transgene was clearly active in the hind stomach and pancreatic epithelium (Fig. 3A-C and 4C, E). In *Gata4-G4Gp-Cre^{TG0}; Rosa26^{lacZ/+}* embryos, β-galactosidase activity was also observed in the stomach epithelium and pancreatic bud, but its expression was broader than the expression of the *Gata4-G4-lacZ* transgene and was present in the epithelium of other endoderm derivatives, including the liver, and duodenum (Fig. 4D, F), indicating that these structures are derived from a *Gata4-G4-lacZ*-expressing progenitor population.

At E13.5, the Cre-dependent *lacZ* expression of the *Gata4-G4Gp-Cre^{TG0}; Rosa26^{lacZ/+}* embryos was detected throughout the endodermal organs of the gastrointestinal and respiratory tract, including fore- and hind stomach, esophagus, duodenum, branching...
epithelium of the lung and in both dorsal and ventral pancreases (Fig. 4H, J). By comparison, the direct activity of the Gata4 G4 enhancer (as measured by Gata4-G4-lacZ expression) was detectable only in a few cells of the dorsal pancreas and hind stomach and was completely absent from other endodermal organs, including the lungs and the liver (Fig. 4G, I, K). The β-galactosidase activity in Gata4-G4Gp-Cre<sup>Tg0</sup>, Rosa26<sup>lacZ+/+</sup> embryos remained robust at later stages of fetal development and after birth in all endodermal organs, including the liver and the lungs (Figs. 4L and 5F), indicating that the Gata4 G4-marked progenitor population stably contributed the majority of cells to all of the endodermal derivatives in the mouse. The specificity of the Gata4 G4 enhancer for the endoderm was clearly evident based on the complete absence of contribution of any Gata4-G4Gp-Cre-marked cells in any organs of mesodermal origin, such as kidney, spleen, and heart (Fig. 5B, D, F). Likewise, Cre activity was completely absent in the mesenchymal compartment of the lung (Fig. 4J). Taken together, these results show that the Gata4 G4 enhancer is expressed in a progenitor cell population, which gives rise to the majority of cells in endoderm-derived organs.

The Forkhead transcription factor FoxA2 binds to a conserved, consensus Forkhead binding site in the Gata4 G4 endoderm enhancer

To identify how Gata4 expression is controlled in the early endoderm via the Gata4 G4 enhancer, we analyzed the Gata4 G4 element for evolutionarily conserved sequences that might serve as potential cis-regulatory elements. These analyses identified a consensus binding site for Forkhead proteins (Fig. 6A). The presence of a perfect, conserved potential binding site for Forkhead transcription factors, combined with the overlapping expression patterns of Gata4-G4-lacZ and foxa2 throughout development (Fig. 2), suggested a potential direct regulatory relationship between FoxA2 and Gata4 via the Gata4 G4 enhancer through the conserved Forkhead site in the enhancer. As an initial test of this hypothesis, we examined the ability of recombinant FoxA2 protein to bind to the Forkhead site in the Gata4 G4 enhancer by Electrophoretic Mobility Shift Assay (EMSA) (Fig. 6B). FoxA2 efficiently bound to the Gata4 G4 Fox site (Fig. 6B, lane 8). Binding was specific because it was competed by excess unlabeled control Forkhead site site (Fig. 6B, lane 9) and by excess unlabeled Gata4 G4 Forkhead site (Fig. 6B, lane10), but not by excess unlabeled mutant versions of either Forkhead site (Fig. 6B, lanes 11, 12). To examine the binding of FoxA2 to the Gata4 G4 enhancer in more detail, we determined the ability of the Gata4 G4 Forkhead site and the mutant version of this site to compete for FoxA2 binding to a bona fide control Fox site. Binding of FoxA2 to the control site was abolished by the addition of unlabeled Gata4 G4 Forkhead site but was not affected by the addition of an excess of mutant version of the Gata4 G4 Forkhead site (Fig. 6B, lanes 4, 6).

To investigate whether the Gata4 G4 Forkhead site is occupied by FoxA2 in cells of endodermal origin, we performed chromatin immunoprecipitation (ChIP) assays in cells of the AR42J pancreatic cell line (Fig. 6C). The AR42J cell line was derived from a chemically-induced rat exocrine pancreatic tumor and expresses numerous endodermal pancreatic transcription factors, including GATA4 and FoxA2 (Aldibbiat et al., 2008;Decker et al., 2006;Rosewicz et al., 1992). Importantly, the DNA fragment encompassing the Forkhead site in the endogenous Gata4 G4 enhancer was precipitated and amplified by PCR when anti-FoxA2 antibody was added to the reaction but not when non-specific IgG was used for ChIP (Fig. 6C). These results demonstrate that endogenous FoxA2 binds to the Forkhead site in the endogenous Gata4 G4 enhancer in its normal chromatin context in endodermally-derived cells. Taken together with the results of the EMSA shown in Fig. 6B, these results establish the Forkhead site in the Gata4 G4 enhancer as a FoxA2 binding site, supporting the notion that Gata4 is a direct transcriptional target of FoxA2 via the Gata4 G4
enhancer. Direct regulation of Gata4 by FoxA2 would be consistent with the overlapping expression of Gata4-G4-lacZ and foxa2 (Fig. 2).

**Gata4 G4 endoderm enhancer activity is dependent on Forkhead-binding site in vivo**

To determine the requirement for Forkhead transcription factor binding to the Gata4 G4 enhancer for function in vivo, we introduced mutations identical to those used in EMSA to completely abolish FoxA2 binding to the site in vitro (Fig. 6B). We then generated transgenic mice harboring the Gata4-G4-lacZ mutant version and determined the effect of the Fox site mutation on enhancer activity compared to the wild type Gata4-G4-lacZ construct at E8.5 (Fig. 7). As in the results shown in Fig. 2, robust expression of the wild type Gata4-G4-lacZ transgene was observed in the definitive endoderm (Fig. 7A, C). In sharp contrast, mutation of the Forkhead binding site in the Gata4 G4 enhancer completely eliminated transgene expression in all three independently generated transgenic lines examined (Fig. 7B, D), indicating a critical role for Forkhead transcription factor binding, likely FoxA2, in the activation of the enhancer. Taken together, the coexpression data showing near precise overlap of the expression of foxa2 and Gata4-G4-lacZ (Fig. 2), the strong binding of FoxA2 protein to the perfect consensus Forkhead site in the enhancer (Fig. 6), and the absolute requirement of the Fox site for enhancer function in vivo, strongly support a pathway in which FoxA2 is a direct transcriptional activator of Gata4 in the early endoderm via the enhancer identified and described in these studies.

**Discussion**

**Dynamic expression and modular regulation of Gata4 during endoderm development**

The Gata4 G4 enhancer described here directs robust expression to the endoderm in the early developing mouse embryo. The activity of this enhancer becomes restricted at midgestation to a small subset of endodermal derivatives in the dorsal foregut and midgut that will form the dorsal pancreas and stomach. Our Cre-based fate mapping experiments show that the total temporal and spatial activity of the Gata4 G4 enhancer from its initial activation broadly in the endoderm at early times to its more restricted activity later results in the marking of all definitive endodermal derivatives from both ventral and dorsal foregut, midgut, and hindgut. When this broad contribution of Cre-marked cells is compared to the very restricted expression of the Gata4-G4-lacZ transgene at later stages of development, the results suggest that the Gata4-G4 positive cells represent a pool of early multipotent endodermal progenitors, and while the activity of the enhancer itself is extinguished, the cells marked by the enhancer at early stages ultimately contribute to essentially all endodermally-derived cells in the embryo. However, we cannot formally exclude a very transient expression of the Gata4-G4 enhancer in other gut domains at later stages that may account for the further contribution of Gata4-G4 positive cells to all endodermal derivatives.

We have recently described another Gata4 endoderm enhancer, referred to as Gata4 G8 (Rojas et al., 2009). Interestingly, the Gata4 G8 enhancer directs expression to the visceral and definitive endoderm at later stages of development than the Gata4 G4 enhancer described in the present studies. The previously described Gata4 G8 enhancer is active in the primitive gut tube and yolk sac beginning at E9.5. In contrast to the Gata4 G4 enhancer described here, the previously described Gata4 G8 enhancer remains active at later stages of development, although its activity also becomes restricted to pancreas, stomach and duodenum and yolk sac (Rojas et al., 2009).

The existence of two endoderm-specific enhancers in the Gata4 gene might be related to the multiple roles of this transcription factor in endoderm development and, more specifically, in the development of pancreas and stomach. Alternatively, the two enhancers might
represent temporally discrete transcriptional units with the early Gata4 G4 enhancer serving as an initiation element and the later Gata4 G8 enhancer serving as a maintenance unit. This type of model would be consistent with previously described examples of transcriptional control in the endoderm and other lineages where multiple enhancers differentially regulate initial activation versus maintenance of gene expression (Chen and Goldhamer, 2004; Meredith et al., 2009; Sasaki and Hogan, 1996; Teboul et al., 2002). It is interesting that the activities of Gata4 G4 and Gata4 G8 enhancers overlap at E9.5 in the dorsal foregut and midgut, a key period in endoderm development when organ buds start to form. The selective advantage of the modularity is not well understood, but perhaps it serves to fine tune the spatiotemporal and quantitative control of Gata4 expression and the subsequent transcriptional programs downstream of this key transcription factor gene.

The rapid and progressive restriction of the Gata4 G4 enhancer activity from E7.5 to E11.5 suggests that the activators of the enhancer might also become restricted during development. Alternatively, the enhancer may contain specific sites for repressors that inhibit enhancer activation in other tissues. One of the most intriguing aspects of the expression pattern directed by the Gata4 G4 enhancer is the exclusive localization to the caudal foregut and midgut region of the gut tube at later times in its activity. It is well known that the developing endoderm displays clear anterior-posterior patterning influenced by signals from the adjacent germ layers that act in a graded manner (Kumar et al., 2003). These signals are critical for the proper development of the gut tube. The specific expression pattern directed by the Gata4 G4 enhancer suggests that it may respond to different thresholds of mesodermal signals that restrict its activity in the precise anterior-posterior pattern observed in the present studies.

**A transcriptional regulatory network for endoderm development by GATA4 and FoxA2**

GATA and Forkhead transcription factors are involved in the specification and differentiation of different endodermal cell types in different organisms, including the fly, the nematode, the mouse, and the fish (Dufort et al., 1998; Fukushige et al., 1998; Mango et al., 1994; Narita et al., 1997; Reiter et al., 2001; Weigel et al., 1989; Zhu et al., 1997). GATA4 and FoxA2 are among the earliest transcription factors to be expressed in the embryonic endoderm in the mouse, and they synergistically activate the expression of endoderm-specific genes (Cirillo et al., 2002; Denson et al., 2000). In this study, we demonstrate that the Gata4 G4 endoderm enhancer requires a conserved Forkhead binding site for function in vivo. Based on the overlapping expression pattern of foxa2 and Gata4-G4-lacZ transgene and the efficient binding of FoxA2 to the Forkhead site in the Gata4 G4 enhancer in vitro and in vivo (Figs. 2, 5), our studies strongly support a role for FoxA2 as a direct transcriptional regulator of Gata4 in the early endoderm.

It would be interesting to determine if the direct regulatory relationship between GATA4 and FoxA2 is reciprocal, since it has been shown in other organisms that GATA factors are likely direct regulators of Forkhead genes. For example, the Drosophila GATA4 ortholog Serpent influences the expression of forkhead in the midgut (Casanova, 1990; Reuter, 1994). Similarly, the promoter of PHA-4, the C. elegans FoxA2 ortholog, is bound and transactivated by two different GATA factors, ELT-1 and ELT-2, in the worm (Azzaria et al., 1996; Horner et al., 1998; Kalb et al., 1998). The reciprocal activation of Gata4 and FoxA2 could serve to reinforce in a feed-forward fashion a regulatory circuit designated to amplify a transcriptional response for specification and differentiation of different endodermal organs.

A mechanism by which FoxA2 and GATA4 might function together in multipotent endoderm cells has been described previously. GATA4 and FoxA2 serve as pioneer factors for the induction of endoderm gene expression, and they were shown to bind to the
chromatin of the silent \textit{Alb} gene in the endoderm prior to the hepatic induction \cite{Bossard1998, Cirillo1998, Zaret1999}. Upon inductive and permissive signals, the two transcription factors relax the compacted chromatin, allowing the binding of other transcription factors to activate the transcription of \textit{Alb} and other genes to initiate the liver program \cite{Bossard1998, Cirillo2002}. Our studies, described here, support a feed forward model for endoderm specification in which FoxA2 directly activates the expression of its own transcriptional partner GATA4. It will be interesting to determine whether this mechanism involving GATA4 and FoxA2 represents a general mechanism by which endoderm competence is predetermined. Genome-wide analyses of FoxA2 and GATA4 bound to their targets in multipotent endodermal cells will provide new insights into how cell specification is achieved along the gut tube.

**Acknowledgments**

We thank Dr. Benoit Gauthier and Dr. Karim Hmadcha for helpful comments on the manuscript. We also thank Jillian Jarrett and Antonio Cárdenas for assistance with these studies. This work was supported by funds from the Instituto Salud Carlos III (CP07/00120, PI080018), Consejería Salud Junta de Andalucía (PI0008) and CIBERDEM to AR and by NIH grants HL099707, HL064658, and DE019118 to BLB.

**References**


Bosse T, Piaseckyj CM, Burghard E, Fialkovich JJ, Rajagopal S, Pu WT, Krasinski SD. Gata4 is essential for the maintenance of jejunal-ileal identities in the adult mouse small intestine. Mol Cell Biol. 2006; 26:9060–70. [PubMed: 16940177]


Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002; 9:279–89. [PubMed: 11864602]


Dev Biol. Author manuscript; available in PMC 2011 October 15.


Fig. 1.
A schematic representation of the mouse *Gata4* locus and the *Gata4*-G4-*lacZ* transgene. The mouse *Gata4* locus, including the seven *Gata4* exons (vertical lines) and the transcriptional start site (bent arrow) is depicted in the top schematic. Conserved non-coding regions among the human, mouse and opossum *Gata4* sequences are represented as red boxes (G1–G9). The lower schematic depicts the *Gata4* G4 intronic sequences cloned into the reporter vector HSP68-*lacZ*. 
Fig. 2.
The Gata4-G4-lacZ transgene is expressed in the visceral and definitive endoderm, dorsal foregut, midgut, and derivatives of the definitive endoderm in the developing stomach and pancreas during mouse embryonic development. Whole mount (A, B, E), sagittal (C, F) and transverse (D, H, K) sections of X-gal stained Gata4-G4-lacZ transgenic embryos are shown. For comparison, whole mount (G) and transverse (I, L) sections of foxa2 in situ hybridization and transverse sections (J, M) of GATA4 immunohistochemistry are shown. White arrowheads mark the definitive endoderm. Transgene expression was only detected transiently in the visceral endoderm of the yolk sac (ys) at E7.5 (A). At slightly later stages, Gata4-G4-lacZ expression appeared in the definitive endoderm (B–D). By E9.5, lacZ expression became restricted to the caudal foregut (fg) and midgut (mg) where expression mirrored the expression of the endodermal marker foxa2 (G, I, L) and endogenous GATA4 (J, M) but was absent in the anterior foregut (black arrows). Bars in all panels=100μm. h, heart; st, septum transversum.
Fig. 3.
Expression directed by the Gata4-G4-lacZ transgene becomes restricted to the stomach epithelium and dorsal pancreatic bud. Whole mount (A) and sagittal sections (B–F) of X-gal-stained Gata4-G4-lacZ transgenic embryos are shown. At E11.5, X-gal staining was observed in the nascent epithelium of the hind stomach and in the pancreatic epithelium (A–C). By E13.5, transgene activity began to diminish in both the hind stomach (arrowhead in D) and the pancreas epithelium (pan epi) (E). No transgene expression was observed at E16.5 (F). Arrow marks the anterior stomach and asterik indicates the stomach mesenchyme in D. (sto) stomach; (eso) esophagus; (duo) duodenum; (pan) pancreas; (pan m) pancreas mesenchyme (li), liver.
The *Gata4* G4 enhancer is active in all endodermal progenitors in the mouse embryo. Yolk sac (A, B), sagittal (C, D) and transverse (E, F) sections and dissected organs (G–L) collected from either *Gata4-G4-lacZ* (A, C, E, G, I, K) or *Gata4-G4Gp-CreTg0*, *Rosa26 lacZ/+* (B, D, F, H, J, L) transgenic embryos were stained for β-galactosidase activity with X-gal. At E11.5, *Gata4-G4-lacZ* expression was absent in the yolk sac and detected only in the epithelium of the hind stomach and pancreas (A, C, E). *Gata4-G4Gp-CreTg0*, *Rosa26 lacZ/+* embryos exhibited a similar pattern of lacZ expression in these tissues, but expression was also apparent in other endoderm-derived including the yolk sac, organs, such as liver, duodenum (B, D, F), indicating that the *Gata4* G4 enhancer was once active in
the progenitors of these organs at earlier stages. By E13.5, the activity of \textit{Gata4-G4-lacZ} transgene activity became weaker and staining was only observed in the dorsal pancreas and hind-stomach (E). No staining was observed in the lungs of \textit{Gata4-G4-lacZ} transgenic embryos (I) at E13.5 or in the liver at E16.5 (K). By contrast, \textit{Gata4-G4Gp-Cre}\textsuperscript{Tg/0}, \textit{Rosa26}\textsuperscript{lacZ/+} embryos exhibited X-gal staining in both dorsal and ventral pancreases, fore- and hind stomach, duodenum, esophagus, liver, and lungs (H, J, L), indicating that these structures were derived from an earlier \textit{Gata4-G4Gp-Cre}-expressing population of cells. Note the specific X-gal staining of \textit{Gata4-G4Gp-Cre}\textsuperscript{Tg/0}, \textit{Rosa26}\textsuperscript{lacZ/+} embryos in the endoderm compartment of the lung (en) but not in the mesenchymal (m) component (H).

(sto), stomach; (pan), pancreas; (dpan), dorsal pancreas; (vpan), ventral pancreas; (h), heart; (nt), neural tube; (li), liver; (lu), lungs; (duo), duodenum.
The Gata4 G4 enhancer is not active in the mesoderm or its derivatives at any stage in development. Representative organs dissected from Gata4-G4-lacZ (A, C, E) and Gata4-G4Gp-CreTg0, Rosa26 lacZ+/ (B, D, F) newborn mice (P0) were assayed for β-galactosidase activity. Gata4-G4Gp-Cre indelibly labeled cells in the lung epithelium of newborn mice (arrowhead in F), consistent with its activity in the endoderm. At that stage, no direct expression of the Gata4-G4-lacZ transgene was observed in the lungs (arrowhead E). No mesodermal organs, including kidney (k) (A, B), spleen (sp) (C, D), or heart (asterisks in E, F) were labeled in either Gata4-G4-lacZ or Gata4-G4Gp-CreTg0, Rosa26RlacZ+/ mice.
Fig. 6.  
*The Gata4-G4-lacZ* transgene contains an evolutionary conserved Forkhead binding site that is bound by FoxA2 protein *in vitro* and *in vivo*. (A) ClustalW analysis of *Gata4* G4 sequences from mouse, human, and opossums revealed a highly conserved, perfect consensus Forkhead binding site (orange box). Asterisks denote nucleotides that have been perfectly conserved among the three species. (B) Recombinant FoxA2 protein was used in EMSA with radiolabeled double-stranded oligonucleotide probe representing a canonical Forkhead binding site from the *Gata4* G2 lateral mesoderm enhancer (Rojas et al., 2005) (lanes 1–6) or the *Gata4* G4 Forkhead site (lanes 7–12). Lanes 1 and 7 contain reticulocyte lysate without recombinant FoxA2 protein. FoxA2 efficiently bound to the Forkhead control site and to the *Gata4* G4 Forkhead site (lanes 2, 8). In both cases, binding was competed by excess, unlabeled control Forkhead site (C, lanes 3 and 9) or unlabeled *Gata4* G4 Forkhead site (G4, lanes 4 and 10), but not by excess unlabeled mutant control (mC, lanes 5 and 11) or an unlabeled mutant version of the *Gata4* G4 (M, lanes 6 and 12) Forkhead sites. (C) FoxA2 binds to the endogenous *Gata4* G4 enhancer *in vivo*. AR42J pancreatic cells were subjected to ChIP to detect endogenous FoxA2 bound to the *Gata4* G4 enhancer using anti-FoxA2 antibody. Following ChIP, the *Gata4* G4 Forkhead site was detected using specific primers flanking the *Gata4* G4 Fox binding site. PCR products were analyzed by agarose gel electrophoresis. Lane 2 contains PCR product obtained following ChIP using anti-FoxA2 antibody (α-Fox). Lane 3 contains PCR product obtained following ChIP using a non-
specific anti-IgG (α-IgG). Lane 1 contains PCR product from input (Inp) amplified prior to immunoprecipitation. Sizes in bp are shown at the left.
Fig. 7.
The *Gata4* G4 early endoderm enhancer is dependent on its conserved Forkhead (Fox) site for activity *in vivo*. Wild type and mutant versions of the *Gata4-G4-lacZ* transgene were used to generate transgenic embryos. Lateral (A, B) and frontal view (C, D) of X-gal stained transgenic embryos at E8.5 are shown. The wild type construct reproducibly directed expression to the definitive endoderm (A, C). Mutation of the *Gata4* G4 Forkhead site completely abolished transgene expression in each of three independently-generated F₀ transgenic embryos analyzed (B, D). Arrowheads mark the definitive endoderm in all four panels.