Regulation of the Fibroblast Growth Factor Receptor 3 Promoter and Intron I Enhancer by Sp1 Family Transcription Factors

(Received for publication, September 23, 1997, and in revised form, November 13, 1997)

Donald G. McEwen and David M. Ornitz‡

From the Department of Molecular Biology and Pharmacology, Washington University Medical School, St. Louis, Missouri 63110

Fibroblast growth factor receptor 3 (FGFR3) has a complex spatial and temporal pattern of expression and is essential for the normal development of a diverse set of tissues. Recently, mutations have been identified in FGFR3 that result in constitutive tyrosine kinase activity and cause a number of different human skeletal disorders. To examine the regulatory mechanisms governing FGFR3 expression, the promoter for the FGFR3 gene was identified and characterized. It resides in a CpG island, which encompasses the 5′ end of the FGFR3 gene and lacks classical cis-regulatory motifs. As little as 100 base pairs of sequence 5′ to the initiation site can confer a 20–40-fold increase in transcriptional activity upon a promoter-less vector. The transcriptional activity of these cis-regulatory sequences is further stimulated by elements found within the first intron. Mapping of the enhancer activity found within intron I identified two purine-rich sequence motifs between +340 and +385. Electrophoretic mobility shift assays demonstrated that sequences within this region bind members of the Sp1 family of transcription factors. In a background lacking Sp1-like activity, we demonstrate that Sp1 can enhance transcription of the minimal promoter (which contains three classical Sp1 sites), whereas both Sp1 and Sp3 can enhance transcription through the elements found in intron I. Although these transcription factors are ubiquitously expressed, we demonstrate that the sequences between −220 and +609 of the FGFR3 gene are sufficient to promote the tissue-specific expression of a reporter gene in transgenic mice.

Cellular differentiation requires the proper interpretation of external stimuli. Although many types of stimuli influence cell fate, a target cell must express the appropriate complement of receptors to perceive, interpret, and respond to environmental signals. The fibroblast growth factors (FGFs) are small molecular mass polypeptides (18–27 kDa), which have been implicated in many developmentally regulated events such as mesoderm induction, angiogenesis, chondrogenesis, malignant transformation, and neuronal differentiation (1, 2). To date, 15 FGF ligands have been described (3, 4). These FGFs all have unique patterns of expression as well as a high affinity for heparin and heparan sulfate proteoglycans (HSPGs). HSPGs have been shown to regulate the biological activity of many FGFs and serve as an essential cofactor required for FGF-induced receptor autophosphorylation (5). Because HSPGs are a major component of the extracellular matrix, they have been implicated in limiting the diffusion of FGFs from their site of production (6). Thus, the spatial and temporal regulation of expression of FGFs provides one mechanism through which FGF-mediated signaling can be regulated during development.

The fibroblast growth factor receptor (FGFR) family consists of four genes, each of which encodes a membrane-spanning tyrosine kinase receptor (3, 7). Recently, both gain-of-function and loss-of-function mutations in the FGFRs have revealed unique roles for these receptors during development (8–13). Specifically, point mutations in FGFR3 have been genetically linked to achondroplasia, thanatophoric dysplasia, and hypochondroplasia (8); all are diseases where bones fail to grow to normal lengths. These skeletal disorders result from defects in the epiphyseal growth plate, a place where FGFR3 is known to be highly expressed (14). When the mutations corresponding to those of achondroplasia (G380R) and thanatophoric dysplasia (R248C and K650E) are introduced into the murine FGFR3 cDNA, ligand-independent activation of the receptor is observed (15, 16). The constitutive activity of this receptor is thought to disrupt normal development by initiating intracellular signals in the absence of ligand. In contrast, loss-of-function alleles of FGFR3 lead to the overgrowth of long bones (10, 12), as well as deafness due to defects in the development of the organ of Corti (10). Although redundancy in the FGFR family may compensate for the loss of FGFR3 activity in some tissues, these results demonstrate that both the regulation of FGFR3 expression and kinase-mediated signaling activity are required for proper development.

To understand the mechanisms that regulate the expression of FGFR3, we have identified and characterized the FGFR3 promoter both in vitro and in vivo. Here we demonstrate that sequences derived from the CpG island found at the 5′ end of the murine FGFR3 gene are capable of promoting transcription in transient transfection assays. Furthermore, the activity of these sequences can be further stimulated by sequences found within the first intron. Localization of the intron enhancer element identified binding sites for the Sp1 family of transcription factors. Characterization of the trans-acting factors demonstrated that Sp1 and Sp3 could promote transcriptional activity through these elements in the Drosophila SL2 cell line. Although Sp1 and Sp3 transcription factors are ubiquitously expressed, the defined minimal promoter and intron enhancer are sufficient to promote the tissue-specific expression of a reporter gene in transgenic mice.

This paper is available on line at http://www.jbc.org
EXPERIMENTAL PROCEDURES

FGFR3 Gene Regulation—FGFR3 expression was determined by RNA protection analysis with the following modifications. Each RNA sample was hybridized with a mix of $5 \times 10^{10}$ cpm FGF3, $5 \times 10^{10}$ cpm $\beta$-gal, and $0.5 \times 10^{10}$ cpm rpL32 RNA probes. Single-stranded RNA probes, as well as duplex RNA hybrids, were purified with RNAzol B (Teltest, Inc.) as described by the manufacturer. The expected sizes for the protected fragments are as follows: FGF3, 430 nt; $\beta$-gal, 288 nt; and rpL32, 179 nt.

Cell Culture and Transfection—NBT-II, CFK2, and RCJ (clone 3,1,C5,18) cell lines were obtained from J. Thiery, J. Henderson, and J. Aubin, respectively. FGF expression was determined as described (23). All cells were transfected in triplicate using a modified calcium phosphate precipitate (15). During each experiment, 5 $\mu$g of each DNA construct (double banded on a CaCl$_2$ gradient) was co-precipitated with 0.5 $\mu$g of Cs$\beta$topo (24). After 24–60 hrs post-transfection, adherent cells were washed with 1 ml of 1 x phosphate-buffered saline, and subsequently assayed for luciferase activity as described (26). $\beta$-Galactosidase activity was determined with the Galacto-Light Plus system, as described by the manufacturer (Tropix, Inc.). To normalize for transfection efficiencies, luciferase values from each transfection were divided by the corresponding $\beta$-galactosidase activities.

The non-adherent $Drosophila$ cell line, SL2, was transfected as described above; however, cells were left in the precipitant-containing media until harvesting at 48 h post-transfection, and they were collected by centrifugation prior to lysis.

Nuclear Extracts and Electrophoretic Mobility Shift Assays (EMSAs)—Nuclear extracts were prepared according to the protocol of Dignam et al. (27). Protein concentrations of the extracts were determined with a modified Bradford assay (Bio-Rad, Inc.). Finally, aliquots of the nuclear extracts were transferred to silanized tubes, snap-frozen in liquid nitrogen, and stored at $-80^\circ$C.

Electrophoretic mobility shift assays were conducted essentially as described (26). Briefly, hybridizations were performed at room temperature for 20 min in a volume of 20 $\mu$l, which consisted of: 0.1 $\mu$mol of protected radioactive activity, 50 $\mu$M 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3. Amplified products were cloned into the HindIII sites of cDNA extracted from the adipose tissue and digested with the same endonucleases.

Heterologous Promoter Plasmids—pRSV(ii)- and p(i)RSV-luc were cloned into the HindIII sites of the plasmid pBluescript KS (Strategene, Inc.). RNase Protection Analysis—Gene expression was determined by RNase protection analysis (22) with the following modifications. Each RNA sample was hybridized with a mix of $5 \times 10^{10}$ cpm FGF3, $5 \times 10^{10}$ cpm $\beta$-gal, and $0.5 \times 10^{10}$ cpm rpL32 RNA probes. Single-stranded RNA probes, as well as duplex RNA hybrids, were purified with RNAzol B (Teltest, Inc.) as described by the manufacturer. The expected sizes for the protected fragments are as follows: FGF3, 430 nt; $\beta$-gal, 288 nt; and rpL32, 179 nt.

Cell Culture and Transfection—NBT-II, CFK2, and RCJ (clone 3,1,C5,18) cell lines were obtained from J. Thiery, J. Henderson, and J. Aubin, respectively. FGF expression was determined as described (23). All cells were transfected in triplicate using a modified calcium phosphate precipitate (15). During each experiment, 5 $\mu$g of each DNA construct (double banded on a CaCl$_2$ gradient) was co-precipitated with 0.5 $\mu$g of Cs$\beta$topo (24). After 24–60 hrs post-transfection, adherent cells were washed with 1 ml of 1 x phosphate-buffered saline, and subsequently assayed for luciferase activity as described (26). $\beta$-Galactosidase activity was determined with the Galacto-Light Plus system, as described by the manufacturer (Tropix, Inc.). To normalize for transfection efficiencies, luciferase values from each transfection were divided by the corresponding $\beta$-galactosidase activities.

The non-adherent $Drosophila$ cell line, SL2, was transfected as described above; however, cells were left in the precipitant-containing media until harvesting at 48 h post-transfection, and they were collected by centrifugation prior to lysis.

Nuclear Extracts and Electrophoretic Mobility Shift Assays (EMSAs)—Nuclear extracts were prepared according to the protocol of Dignam et al. (27). Protein concentrations of the extracts were determined with a modified Bradford assay (Bio-Rad, Inc.). Finally, aliquots of the nuclear extracts were transferred to silanized tubes, snap-frozen in liquid nitrogen, and stored at $-80^\circ$C.

Electrophoretic mobility shift assays were conducted essentially as described (26). Briefly, hybridizations were performed at room temperature for 20 min in a volume of 20 $\mu$l, which consisted of: 0.1 $\mu$mol of protected radioactive activity, 50 $\mu$M 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3. Amplified products were cloned into the HindIII sites of the plasmid pBluescript KS (Strategene, Inc.). RNase Protection Analysis—Gene expression was determined by RNase protection analysis (22) with the following modifications. Each RNA sample was hybridized with a mix of $5 \times 10^{10}$ cpm FGF3, $5 \times 10^{10}$ cpm $\beta$-gal, and $0.5 \times 10^{10}$ cpm rpL32 RNA probes. Single-stranded RNA probes, as well as duplex RNA hybrids, were purified with RNAzol B (Teltest, Inc.) as described by the manufacturer. The expected sizes for the protected fragments are as follows: FGF3, 430 nt; $\beta$-gal, 288 nt; and rpL32, 179 nt.

Cell Culture and Transfection—NBT-II, CFK2, and RCJ (clone 3,1,C5,18) cell lines were obtained from J. Thiery, J. Henderson, and J. Aubin, respectively. FGF expression was determined as described (23). All cells were transfected in triplicate using a modified calcium phosphate precipitate (15). During each experiment, 5 $\mu$g of each DNA construct (double banded on a CaCl$_2$ gradient) was co-precipitated with 0.5 $\mu$g of Cs$\beta$topo (24). After 24–60 hrs post-transfection, adherent cells were washed with 1 ml of 1 x phosphate-buffered saline, and subsequently assayed for luciferase activity as described (26). $\beta$-Galactosidase activity was determined with the Galacto-Light Plus system, as described by the manufacturer (Tropix, Inc.). To normalize for transfection efficiencies, luciferase values from each transfection were divided by the corresponding $\beta$-galactosidase activities.

The non-adherent $Drosophila$ cell line, SL2, was transfected as described above; however, cells were left in the precipitant-containing media until harvesting at 48 h post-transfection, and they were collected by centrifugation prior to lysis.

Nuclear Extracts and Electrophoretic Mobility Shift Assays (EMSAs)—Nuclear extracts were prepared according to the protocol of Dignam et al. (27). Protein concentrations of the extracts were determined with a modified Bradford assay (Bio-Rad, Inc.). Finally, aliquots of the nuclear extracts were transferred to silanized tubes, snap-frozen in liquid nitrogen, and stored at $-80^\circ$C.

Electrophoretic mobility shift assays were conducted essentially as described (26). Briefly, hybridizations were performed at room temperature for 20 min in a volume of 20 $\mu$l, which consisted of: 0.1 $\mu$mol of protected radioactive activity, 50 $\mu$M 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3; 5'-GTC GAC GCG TGT GAG TTG GGC TCT TTG CCT CGC AG-3. Amplified products were cloned into the HindIII sites of the plasmid pBluescript KS (Strategene, Inc.). RNase Protection Analysis—Gene expression was determined by RNase protection analysis (22) with the following modifications. Each RNA sample was hybridized with a mix of $5 \times 10^{10}$ cpm FGF3, $5 \times 10^{10}$ cpm $\beta$-gal, and $0.5 \times 10^{10}$ cpm rpL32 RNA probes. Single-stranded RNA probes, as well as duplex RNA hybrids, were purified with RNAzol B (Teltest, Inc.) as described by the manufacturer. The expected sizes for the protected fragments are as follows: FGF3, 430 nt; $\beta$-gal, 288 nt; and rpL32, 179 nt.
pronounce of FVB/N embryos, and founder animals were identified as described previously (30). Founder animals were bred to wild type FVB/N mice, and the resulting offspring were analyzed for expression of the transgene. With the exception of p(2951/+609)/FR3 luc2 construct, multiple independent transgenic lines were analyzed to control for insertion site effects on transgene expression.

RESULTS

Identification of a CpG Island at the 5’ End of the FGFR3 Gene—Genomic DNA encompassing the 5’ end of the FGFR3 gene was isolated using a cDNA-derived DNA probe. Three kilobases of DNA extending 5’ from the end of the published FGFR3 cDNA and the complete 5’-untranslated region (UTR) were sequenced. Analysis of genomic sequences encompassing the 5’ end of the FGFR3 gene identified a CpG island (Fig. 1A), whereas comparisons with the cDNA sequence (GenBank accession no. M81342) identified a 376-bp intron within the 5’-UTR sequences. The intron/exon organization of the 5’ end of the FGFR3 gene as well as the donor and acceptor splice sites are indicated in Fig. 1 (A and B). To determine the start site of transcription, two fragments (p222a2 and p222b3) corresponding to genomic DNA that overlaps the 5’-UTR sequences, and lowercase denotes intron I. Splice donor and acceptor sites are shown in bold, and consensus Sp1 binding sites are shaded. The start site of transcription is indicated as +1. The translational start site and coding sequence for murine FGFR3 are denoted by double underlining and the single-letter amino acid code. C, determination of transcriptional initiation start site. A RNase protection probe, corresponding to p222b3 (shown in A), was hybridized to mouse brain RNA at 65°C. After digesting single-stranded RNA, the protected product was resolved along with a sequencing ladder, and the start site determined relative to the sequence of the probe.

Therefore, this region of the FGFR3 gene was tested for transcriptional regulatory activity.

Basal Promoter Activity of the FGFR3 CpG Island—To demonstrate the transcriptional potential of the sequences near the 5’ end of the FGFR3 gene, various fragments of proximal sequence were cloned upstream of the firefly luciferase reporter gene (luc). A 5’ deletion series with a fixed 3’ end at the –27 position was generated through restriction endonuclease digestion. The constructs were transfected into a series of different cell lines that either express (NBT-II and CFK2 cell lines) or fail to express (NIH/3T3 and RCJ cell lines) the endogenous FGFR3 gene. All deletion constructs tested were capable of inducing a 20–40-fold increase in luciferase activity from a promoter-less vector (pGL2-Basic) (Fig. 2A). Nevertheless, this activity, at least in transient transfection assays, is independent of cellular background and fails to mimic the expression pattern of the endogenous gene.

With the exception of the p(–2311/–27)/FR3-luc (c2) construct in the CFK2 cells (activity to be described elsewhere),

no significant differences in transcriptional potential were observed when the sequences between –2951 bp and –220 bp were deleted (Fig. 1A). This observation suggests that the most proximal 193 nt of the FGFR3 gene are capable of initiating transcription. To confirm the role of the sequence between –220 and –27 (the proposed minimal promoter), a 3’ terminal deletion that removed these sequences (p(–1537/–27)/FR3-luc, c5) was tested for promoter function in NBT-II and NIH/3T3 cells. In NBT-II cells, the 20-fold increase in transcriptional activity of the p(–1537/–27)/FR3-luc (c3) construct is dependent on the proximal sequences; when the proximal sequences are deleted, promoter activity returns to background levels (Fig. 2B; compare c5 to c3). Identical results were obtained with NIH/3T3 cells (data not shown). These results demonstrate that the sequences between –220 and –27 are

2 D. G. McEwen, M. Naski, D. Towler, and D. M. Ornitz, manuscript in preparation.
FGFR3 Gene Regulation

Fig. 2. Functional analysis of mouse FGFR3 5' cis-regulatory sequences in NBT-II, NIH/3T3, and CFK2 cell lines. A, a panel of 5' terminal deletion-luciferase reporter constructs, p(−2951/−27)FR3-luc (c1), p(−2311/−27)FR3-luc (c2), p(−1537/−27)FR3-luc (c3), and p(−220/−27)FR3-luc (c4), were transfected into various cell lines. The 5' and 3' boundaries of each construct are denoted in parentheses. B, a 3' terminal deletion (p(1537/−220)FR3-luc (c5)) of sequences between −220 and −27 in c3 was tested for transcriptional activity in NBT-II cells. C, assessment of 5' deletions p(−175/−27)FR3-luc (c6), p(−126/−27)FR3-luc (c7) and p(−79/−27)FR3-luc (c8), of c4. All bar and column charts show data representative of at least two independent experiments. Constructs were transfected in triplicate, and cell lysates harvested 60 h post-transfection. Values are plotted as fold induction over the empty vector (pGL2-Basic) and are calculated by dividing the mean (± standard deviation) derived for each construct by the mean of the pGL2-Basic construct (pGL2-Basic = 1).

required for transcriptional initiation.

Five potential Sp1 binding sites are found within the 193-bp proximal promoter region (Fig. 1B). To further examine the transcriptional activity of the minimal promoter, the effects of additional 5' terminal deletions were examined. Deletion of sequences between −220 and −126 resulted in no significant loss of transcriptional activity, whereas a 43%, 43%, and 34% decrease in transcriptional activity was observed in NBT-II (Fig. 2C), NIH/3T3 (data not shown), and CFK2 cells (data not shown), respectively, when the sequences between −126 and −79 were deleted. This loss in activity most likely reflects a loss of the additive effects of multiple Sp1 binding sites on transcription; nevertheless, this 52-bp fragment still induces at least a 15–20-fold increase in reporter gene expression. These observations suggest that a single Sp1 site (or some other undefined site(s)) between −79 and −27 can serve to initiate transcription.

Localization of Enhancer Elements in FGFR3 Intron I—The transcriptional activity of basal promoters is often regulated by enhancer elements found at some distance from the start site of transcription (37). To screen for potential enhancer-like elements, constructs were generated that contained the sequences encoding the 5'UTR and the first intron. Addition of the sequences between 126/−27 can serve to initiate transcription.

Functional analysis of mouse FGFR3 5' cis-regulatory sequences in NBT-II, NIH/3T3, and CFK2 cell lines.

FGFR3 Gene Regulation

In many cases, transcription factor binding sites are modular and can confer transcriptional activity upon heterologous promoters. To determine if the sequences within intron I could influence the transcriptional activity of a heterologous promoter, sequences corresponding to intron I (i, +132 to +508) were positioned both 5' and 3' of the Rous sarcoma virus (RSV) minimal promoter. Placement of the intron 3' to the RSV-minimal promoter resulted in a 15- and 80-fold enhancement of RSV-mediated gene expression in CFK2 and NBT-II cells, respectively (compare pRVS(i)−luc to pRVS-luc, Fig. 4, A and B). Furthermore, the same intron sequences can enhance the activity of the RSV promoter when placed in the 5' position (p(i)RSV-luc; Fig. 4, A and B). However, the degree of stimulation is much less in the 5' position (4- and 11-fold in CFK2 and NBT-II cells, respectively) than in the 3' position. These data, in conjunction with the data shown in Fig. 3, demonstrate that enhancer activity is found within intron I, and that these
sequences can enhance the transcriptional potential of both the homologous and a heterologous promoter.

Mapping Enhancer Activity to Polypurine Stretches in Intron I—Although the intron preferentially enhances transcription from the 3’ position, manipulation of the intron sequences, when in the 3’ position, could affect the efficiency of mRNA splicing, mRNA stability, and/or proper translational initiation of the luciferase reporter. To avoid these potential problems and still identify the cis-acting enhancer elements found in intron I, the transcriptional potential of a series of overlapping 60-bp fragments derived from the intron was assessed in the 5’ position, relative to the RSV minimal promoter. Three overlapping fragments, +310/+368, +339/+395, and +369/+427 (henceforth designated δ, η, and α, respectively), were capable of enhancing reporter activity in both CFK2 (Fig. 5B) and NBT-II cells (data not shown), whereas all other fragments showed little or no stimulatory activity in both cell lines. To further refine this analysis, subdivision of the η fragment into three overlapping 30-mers, designated A, B, and C in Fig. 5A, was conducted. Although all three fragments could enhance the transcriptional activity of the RSV minimal promoter, only fragment B (+354 to +384) could promote transcriptional activity that was comparable to either the complete intron or to the η fragment (Fig. 5C).

To identify other genes that might be regulated by similar sequences, the nucleotide data base (GenBank) was searched with the BLAST sequence analysis tool (40). Several genes were identified that contained sequences similar to two purine-rich direct repeats found within the η and B fragments (Fig. 5A) (41, 42). Notably, the human epidermal growth factor receptor gene (EGFR) contains several purine-rich sequences, which are similar to those found in the B fragment. Johnson et al. (42) demonstrated that the GGAGGAGG element found within the EGFR promoter can bind to the ubiquitously expressed transcription factor Sp1.

Sp1-like Molecules Bind Polypurine Enhancer Elements—To determine the potential role of Sp1 in regulating the expression of FGFR3, EMSAs were utilized to characterize the transacting factors that bind to both the η and B fragments. Three independent DNA–protein complexes were observed when nuclear extracts, prepared from the NBT-II cells, are incubated with the η fragment from intron I (Fig. 6A, lanes 2). Although the slowest migrating complex is competed away by a nonspecific competitor (γ fragment in Fig. 6A, lanes 3–5), two specific DNA–protein complexes were identified (Fig. 6A, lanes 7–10). These complexes were disrupted through competition with as little as 10-fold molar excess of unlabeled probe (Fig. 6A, lanes 8–10).

Fig. 3. Identification of an enhancer activity in intron I of the murine FGFR3 gene. A, the effects of the sequences encoding the entire 5’-UTR and intron I upon luciferase reporter activity were assessed in the NBT-II cell line. B, the role of the endogenous start site of transcription was determined by comparing the transcriptional activity of constructs that contain the entire UTR and intron sequences (c9) to constructs with (c12) or without (c11) the endogenous start site of transcription. C, the ability of the intron to regulate expression was demonstrated by comparing a construct that contains the UTR and intron sequences (c11) to a construct that only contained the UTR sequences (c12) or a construct that lacked the UTR sequences completely (c14). D, the ability of intron I to affect transcription was demonstrated by analyzing the ability of intron-containing, but UTR-lacking constructs (c14) to constructs with (c11) or without (c4) the complete UTR/intron sequences. cis-Regulatory sequences derived from the 5’ end of the murine FGFR3 gene are depicted with horizontal lines, untranslated sequences (found in exons 1 and 2) are represented by open boxes, intron sequences are denoted by zig-zag lines, and luciferase coding sequences are shown as shaded boxes. Map coordinates are shown above each construct and correspond to the sequence shown in Fig. 1A. Data are represented as -fold induction, and were determined as described in the legend for Fig. 2 except that values are plotted relative to the appropriate basal promoter construct, which lacked all sequences between −26 and +609.

Fig. 4. Intron I sequences can regulate the Rous sarcoma virus minimal promoter from both the 3’ and 5’ positions. Constructs that contain intron I sequence (i) from +132 to +508 of the murine FGFR3 gene cloned either 5’ to or 3’ to the RSV minimal promoter were tested for transcription enhancing activity in CFK2 cells (A) and NBT-II cells (B).
FGFR3 Gene Regulation

8–10). Significant competition for the binding of nuclear factors is also observed with the B fragment (Fig. 6A, lanes 11–14) and C fragments (data not shown) of η, whereas the A fragment requires 5–10-fold higher concentrations to compete effectively (data not shown).

Utilizing the B fragment (+354/+384) as a probe, four independent protein-DNA complexes (labeled I–IV) could be resolved with nuclear extracts from the NBT-11 cells (Fig. 6B, lane 2). To examine both the specificity and identity of these complexes, various unlabeled duplex oligonucleotides were used to compete for the binding of nuclear proteins. Poor competition with an unrelated kidney-specific transcription factor was capable of competing for proteins in complexes II, III, and IV, which bind to the sequences between +354 and +384 (Fig. 6B, lanes 9–12). Complex I was not competed away by a 50-fold excess of either the EGFR (Fig. 6B, lane 8), the Sp1 (Fig. 6B, lane 12), or the A fragment (Fig. 6B, lane 16) competitors, and was only lost when DNA-protein complexes were challenged with the unlabeled C fragment (Fig. 6B, lanes 24). These results suggest that Sp1 or Sp1-like molecules are present in bands II, III, and IV, whereas band I does not depend upon Sp1-like DNA binding activity.

The transcriptional role for the various DNA-protein complexes was further examined by incubating the B fragment with extracts from several different cell lines. Examination of the resulting complexes show that CFK2 and RCJ cell lines lack band I (Fig. 6C, lanes 3 and 4). However, the ability of the sequences between +11 and +609 to promote a 60- and 15-fold increase in transcriptional activity in CFK2 and RCJ cells, respectively, suggests that band I is not required for the function of the enhancer element found in intron I (Fig. 3B). Further comparisons between the extracts revealed that band III is the predominant complex found in RCJ cell extracts. As shown in Fig. 3 (B and C), the capacity of the intron sequence to enhance transcription in the RCJ cell line is limited (2–10-fold) in comparison to the CFK2 cell line, thus suggesting that complexes found within either band II, band IV, or both are required for optimal transcriptional activity.

Sp1 Family Members Regulate Transcription of the FGFR3 Promoter in SL2 Cells—To confirm the trans-activation of the promoter sequences by members of the Sp1 family of transcription factors, the Drosophila SL2 cell line was utilized. SL2 cells
FIG. 6. Characterization of nuclear factors binding to elements in intron I. A, DNA binding activity of the η fragment with nuclear extracts isolated from NBT-II cells. EMSAs were performed with radiolabeled η fragment (lanes 2, 7, and 11). The specificity of these DNA-protein complexes was assessed through competition with one nonspecific (χ, lanes 3–5) and two specific competitors (η, lanes 8–10; B fragment, lanes 12–14). B, DNA-protein interaction with the B fragment and nuclear extracts derived from NBT-II cells reveal four independent complexes. The specificity of these complexes was assessed through competition with a series of defined transcription factor binding sites: lanes 2–4, a kidney-specific transcription factor binding site (LFK); lanes 6–8, a binding site derived from the human EGFR (EGFR); lanes 9–12, a classical Sp1 GC box (Sp1); lanes 14–16, fragment A; lanes 18–20, fragment B; lanes 22–24, fragment C. EMSAs of the B fragment with a nuclear extracts obtained from NBT-II (lane 2), CEFK2 (lane 3), and RCJ (lane 4) cells.

lack Sp1-like activity and have been used by others to directly assess the transcriptional role of Sp1 and Sp1-like family members (19, 33, 34, 44). To determine if Sp1 could regulate the transcriptional potential of the FGFR3 promoter, p(−2951/+609)FR3-luc (c1) and p(−2951/+609)/FR3-luc (c9) constructs were co-transfected with an expression vector for Sp1. Expression of Sp1 resulted in a 13- and 25-fold increase in the transcriptional activity of −2951/−27 and −2951/+609 promoter fragments, respectively (Fig. 7A). This regulatory activity is comparable to the 11-fold stimulation seen with the Sp1-responsive dihydrofolate reductase promoter (33) (Fig. 7A). To examine whether Sp1 or Sp3 could regulate expression through the A, B, and C fragments various RSV-luc reporter constructs were transfected into SL2 cells along with expression vectors for either Sp1 or Sp3. Both Sp1 and Sp3 weakly stimulated the activity of pA-RSV-luc and pC-RSV-luc, whereas pB-RSV-luc was stimulated 15-fold and 12-fold by Sp1 and Sp3, respectively. These data, and the ability of Sp1 consensus competitors to compete for binding of nuclear proteins to the intron enhancer elements, demonstrate that members of the Sp1 family of transcription factors can promote transcription of the FGFR3 promoter.

Minimal cis-Regulatory Sequences Confer Tissue-specific Expression in Vivo—The reporter constructs containing various FGFR3 promoter fragments direct transcription in all four cell lines tested, regardless of whether the endogenous FGFR3 gene is expressed (Fig. 2A). However, FGFR3 is expressed in a very cell-specific pattern in vivo (14). To examine the tissue-specific expression pattern of the defined promoter sequences, various reporter constructs were used to generate transgenic animals. RNase protection assays were used to simultaneously monitor the expression of the endogenous FGFR3 gene, the transgene gene (β-gal), and a ubiquitously expressed ribosomal RNA molecule (rpl32). The endogenous gene is detected to some extent in all tissues examined (Fig. 8). However, the reporter gene is only expressed in a subset of these tissues in all lines examined. Comparisons of reporter gene expression between independent transgenic lines derived with increasing lengths of 5′ cis-regulatory sequences demonstrated that the elements found between −220 and +609 are sufficient to direct the tissue-specific expression of a reporter gene in lung, liver, small intestine, kidney, and skin (Fig. 8, compare −220/+609 to −15 (kilobase pairs)/+609). Furthermore, 3 of the 13 lines examined also expressed β-gal mRNA in the brain. However, the inconsistent nature of this expression suggests that it results from insertion site effects. The ability of these proximal sequences to promote the tissue-specific expression of the reporter gene in vivo demonstrates that much of the tissue-specific regulation of the FGFR3 gene is controlled by the regulatory elements found between −220 and +609.

DISCUSSION

The promoter for the FGFR3 gene resides in a CpG island that lacks the classical CAAT box and TATA box motifs found in many eukaryotic promoters. Sequence analysis of the FGFR3 promoter revealed a number of transcription factor binding sites, including five classical Sp1 sites, within the first 200 bp 5′ of the transcription start site. The positioning of the basal transcriptional machinery in a TATA-less promoter can occur independent of InR sequences when Sp1 binding sites are present (44–46). In such instances, Sp1 is capable of stabilizing transcriptional initiation complexes approximately 50 bp downstream from an Sp1 binding site (45). Mapping of the start site of transcription was achieved through RNAase protection, and it was shown that transcriptional initiation occurs 22 bp 5′ from the end of the longest published mouse FGFR3 cDNA (GenBank accession no. M81342) and 57 bp 3′ of the most proximal Sp1 binding site. Our start site differs by only two nucleotides from that previously described by Perez-Castro et al. (47) and both start sites are positioned such that Sp1 could facilitate organization of the transcription initiation complex.
structures that contained the 5′-UTR sequences but lacked intron I failed to result in significant transcriptional enhancement, whereas placement of the intron alone 3′ relative to the FGFR3 promoter sequences afforded the same transcriptional enhancement seen with the UTR/intron combination. These results demonstrated that the enhancer-like activity resides in intron I.

Mapping of the intron enhancer activity to sequences between +340 and +395 identified two polyuridine direct repeat sequence motifs. The sequence and organization of these motifs is similar to a motif previously identified in the EGFR promoter. From the studies of Johnson et al. (42), it was determined that this site was capable of enhancing the transcription of the EGFR promoter in vitro. Through their studies, they also showed that these elements were sensitive to S1 nuclease and bound the Sp1 transcription factor. Although this site in the FGFR3 promoter is not sensitive to S1 nuclease, it does interact with Sp1-like DNA binding activity as shown through gel shift analysis. The specificity of this interaction was demonstrated through competition with the classical Sp1 (GC box) binding site and a non-classical Sp1 binding site derived from the promoter for the human EGFR. The classical Sp1 element is unrelated to the polyuridine stretch; however, it was capable of competing for the DNA binding activity found in all but one of the resulting DNA-protein complexes. Furthermore, transfection studies in SL2 cells demonstrated that Sp1 could promote transcriptional activity through either the basal promoter alone or the promoter/enhancer combination whereas co-transfection studies demonstrated that both Sp1 and Sp3 can enhance transcription through the intron enhancer element. Together, these data suggest that binding sites for members of the Sp1 family of transcription factors, both proximal and distal to the start site of transcription, can work together to enhance transcription of the FGFR3 gene.

The ability of proximal and distal Sp1 binding sites to synergistically regulate transcription has been observed by others (50, 52, 53). Through these studies, it has been shown that Sp1-Sp1 protein interactions can induce looping of the intervening cis-regulatory sequences. These proximal-distal interactions are hypothesized to regulate gene transcription by increasing the local concentration of Sp1 glutamine-rich activation domains near the start site. Such a model would explain the synergistic ability of the intron enhancer to regulate the transcriptional activity of the FGFR3 basal promoter. Although Sp3 has usually been shown to serve a negative regulatory role by competing for Sp1 binding sites, at least two other studies have demonstrated that SP3 can promote transcriptional activity (54, 55). This transcription-promoting ability of Sp3 in our experiments may reflect the sequence-specific context of the binding site, as evidenced by the ability of Sp3 to transactivate the B fragment. Additional experiments will be required to demonstrate the in vivo role of Sp3 in FGFR3 promoter regulation.

Transient transfection assays demonstrated that promoter activity resides in the CpG island found at the 5′ end of the FGFR3 gene, whereas an enhancer element was located in the first intron. However, this activity failed to parallel the cell-type specific expression of the endogenous FGFR3 gene. To assess whether or not the minimal promoter sequences defined above were capable of promoting cell-type specific expression in vivo, various lengths of regulatory sequences were used to generate transgenic animals. Surprisingly, analysis of the transgene expression in 13 independent transgenic lines demonstrated that the sequences between −220 and +609 provide

---

3 D. G. McEwen and D. M. Ornitz, unpublished observations.
the proper regulatory elements required for the expression of a reporter gene in a subset of the tissues that normally express the endogenous FGFR3 gene, whereas these same elements fail to limit the cell-specific expression pattern of the endogenous FGFR3 gene in vitro. These data suggest that other undefined mechanisms exist to regulate the expression of the FGFR3 minimal promoter in vivo.

Due to the increased mutability associated with 5-methylcytosine, the conservation of the CpG island at the 5' end of the FGFR3 gene suggests that it plays some important regulatory role in vivo. One possible way in which these sequences might regulate gene expression in a tissue-specific manner is through the methylation of any of the 83 CpG dinucleotides found within the −220 to +609 region. The establishment of methylation patterns during development (56) is required for embryo viability, and has been shown to regulate the transcriptional activity of many genes by either directly interfering with the binding of transcription factors to their DNA cognates (57–59) or by recruiting methyl binding transcriptional repressor proteins (60, 61). Unlike the hypomethylated state of most CpG islands, preliminary studies in which we examined the methylation status of the FGFR3 CpG island in numerous tissues, as well as the transcriptional activity of in vitro methylated reporter constructs, suggested that methylation may be a contributing factor to the tissue specificity exhibited by the FGFR3 minimal promoter in vivo.

Acknowledgments—We thank I. Boime, J. Gordon, and D. Towler for their insightful discussions. We also thank J. Azizkhan for help in establishing the SL2 cell culture system and C. Neville for help in assembling this manuscript.

REFERENCES
1. Basilio, C., and Moscatelli, D. (1992) Adv. Cancer Res. 59, 115–165
2. Wilkie, A. O. M., Morris-Ray, G. M., Jones, E. Y., and Heath, J. K. (1995) Curr. Biol. 5, 500–507
3. Coulter, P., Pontarotti, P., Rouhini, R., Hartung, H., Goldfarb, M., and Birnbaum, D. (1997) J. Mol. Evol. 44, 43–56
4. McWhirter, J. R., Goulding, M., Weiner, J. A., Chun, J., and Merlie, J. P. (1997) J. Mol. Biol. 269, 248–252
5. Naski, M. C., Wang, Q., Xu, J., and Ornitz, D. M. (1996) Nat. Genet. 13, 233–237
6. Webster, M. K., and Donoghue, D. J. (1996) EMBO J. 15, 520–527
7. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
8. Boudreaux, J. M., and Towler, D. A. (1996) J. Biol. Chem. 271, 7508–7515
9. Courey, A. J., and Tjian, R. (1988) Cell 55, 887–898
10. Hagen, G., Muller, S., Beato, M., and Suske, G. (1984) EMBO J. 3, 3843–3851
11. Shen, M. M., and Leder, P. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 8240–8244
12. Ausubel, F. M. (ed) (1994) in Current Protocols in Molecular Biology. (Janssen, K., ed) John Wiley & Sons, New York
13. McWhirter, J. R., Goulding, M., Weiner, J. A., Chun, J., and Merlie, J. P. (1997) J. Biol. Chem. 269, 248–252
14. Peters, K., Ornitz, D. M., Werner, S., and Williams, L. (1993) J. Biol. Chem. 268, 24892–24901
15. Naski, M. C., Wang, Q., Xu, J., and Ornitz, D. M. (1996) Nat. Genet. 13, 947–954
16. Smale, S. T., and Baltimore, D. (1989) Cell 57, 4878–4884
17. Wang, Q., Xu, J., and Ornitz, D. M. (1996) Nat. Genet. 13, 947–954
18. Shen, M. M., and Leder, P. (1992) Oncogene 7, 308–313
19. Courey, A. J., and Tjian, R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3879–3883
20. Colvin, J. S., Bohne, B. A., Harding, G. W., McEwen, D. G., and Ornitz, D. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5067–5072
21. Shen, M. M., and Leder, P. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 8240–8244
22. Ausubel, F. M. (ed) (1994) in Current Protocols in Molecular Biology. (Janssen, K., ed) John Wiley & Sons, New York
23. McWhirter, J. R., Goulding, M., Weiner, J. A., Chun, J., and Merlie, J. P. (1997) J. Biol. Chem. 269, 248–252
24. Swick, A. G., Blake, M. C., Kahn, J. W., and Azizkhan, J. C. (1989) Nucleic Acids Res. 17, 9291–9304
25. Chen, X., Azizkhan, J. C., and Lee, D. C. (1992) Oncogene 7, 1805–1815
26. Morgan, W. D. (1989) Mol. Cell. Biol. 9, 4099–4104
27. Boisclair, Y. R., Brown, A. L., Casola, S., and Rechler, M. M. (1993) J. Biol. Chem. 268, 24892–24901
28. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475–1486
29. Johnson, S. B., Rothstein, J. L., and Knowles, B. B. (1989) J. Biol. Chem. 264, 177–181
30. Iannello, R. C., Young, J., Sumarsono, S., Tymms, M. J., Dahl, H., and Kola, I. (1997) Mol. Cell. Biol. 17, 1123–1134
31. Connor, A. J., and Tjian, R. (1991) Genes Dev. 5, 1646–1656
32. Su, W., Jackson, S. P., and Tjian, R. (1989) Cell 59, 827–836
33. Colvin, J. S., Bohne, B. A., Harding, G. W., McEwen, D. G., and Ornitz, D. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 83, 5670–5674
34. Morgan, W. D. (1989) Mol. Cell. Biol. 9, 4099–4104
35. Morgan, W. D. (1989) Mol. Cell. Biol. 9, 4099–4104
36. Boisclair, Y. R., Brown, A. L., Casola, S., and Rechler, M. M. (1993) J. Biol. Chem. 268, 24892–24901
37. Li, E., Bestor, T. H., and Jaenisch, R. (1992) Cell 72, 450–456
38. Johnson, A. C., Jinno, Y., and Merlino, G. T. (1988) J. Biol. Chem. 263, 225–232
39. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
40. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
41. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
42. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
43. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
44. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
45. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
46. Liang, Y., Robinson, D. F., Dennig, J., Suske, G., and Fahl, W. E. (1996) J. Biol. Chem. 271, 4174–4181
Regulation of the Fibroblast Growth Factor Receptor 3 Promoter and Intron I Enhancer by Sp1 Family Transcription Factors
Donald G. McEwen and David M. Ornitz

J. Biol. Chem. 1998, 273:5349-5357.
doi: 10.1074/jbc.273.9.5349

Access the most updated version of this article at http://www.jbc.org/content/273/9/5349

Alerts:
  - When this article is cited
  - When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 59 references, 26 of which can be accessed free at http://www.jbc.org/content/273/9/5349.full.html#ref-list-1