Identification of positive and negative regulatory elements involved in the retinoic acid/cAMP induction of Fgf-3 transcription in F9 cells

Akira Murakami, Daniel Grinberg⁺, Jane Thurlow and Clive Dickson*

Imperial Cancer Research Fund Laboratories, PO Box 123, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

Received September 1, 1993; Accepted October 11, 1993

ABSTRACT

The proto-oncogene Fgf-3 has been implicated as an important signalling molecule in vertebrate development. In the mouse, it is expressed for a limited time at a multitude of sites from embryonic day 7 to birth. Transcription of Fgf-3 initiates at three promoter regions resulting in the generation of various mRNAs which nevertheless all encode the same protein products. A 1.7kb DNA fragment which encompasses these regions was joined to the CAT reporter gene and shown to function as a promoter in embryonal carcinoma cells. In stable transfectants the promoter retains its retinoic acid inducibility, initiating transcription at the same cap-sites as the endogenous gene. In differentiated F9 cells, transient transfection of progressive and targeted deletion mutants of the promoter region has revealed at least two positive and three negative regulatory elements. With one exception, loss of these elements was shown to dramatically affect promoter activity in stable transfectants of F9 cells. However the promoter remained inducible by retinoic acid to differing degrees, apart from deletions encompassing PS-4A which essentially abolished promoter activity in both undifferentiated and differentiated cells. The sequences of these potential regulatory regions were further defined using DNase-I footprinting, revealing some similarities to consensus binding sites for known transcription factors.

INTRODUCTION

Fgf-3 is one of eight low molecular weight polypeptides which together form the fibroblast growth factor family [1, 2, 3, 4]. Apart from the well known property of inducing proliferation for a wide variety of cell types, some members of this family have been shown to enhance cell migration, and depending on context, promote or inhibit cell differentiation [1]. Fgf-3 (previously known as int-2), was discovered as a proto-oncogene

transcriptionally activated by proviral insertion in mouse breast tumors [5, 6]. Transcripts of this gene were not detected in the normal mammary gland, nor in most adult mouse tissues, although this does not exclude expression in minor, specialised cell populations. Therefore the inappropriate expression of Fgf-3 was postulated to be instrumental in the induction of breast tumors, a notion which gained considerable support from subsequent studies in transgenic mice, where ectopic expression of the gene led to hyperplastic lesions reminiscent of the virally induced disease [7, 8, 9]. Normal expression of the gene has been revealed by *in situ* hybridization techniques and indicate that the gene is active at specific times in a variety of sites in the developing mouse embryo [10, 11, 12, 13]. Expression can first be detected in the parietal endoderm at embryonic day 7 (E7) and then at several sites including the migrating mesoderm of post-implantation embryos, and at later times in the hindbrain, purkinji cells of the cerebellum, mesenchyme of the tooth bud, the retina and the sensory cells of the inner ear. These patterns of expression suggest multiple roles for Fgf-3, possibly related to proliferation, cell migration or to some inductive capacity. More direct evidence for a developmental function came from gene deletion studies using homologous recombination to inactivate Fgf-3 [14]. In this study a major defect in the posterior region of the fetus was evident, which was manifest as a very short tail. Additionally, there was a defect of the inner ear, although this appeared not to be fully penetrant [14]. *In situ* hybridization studies with *Xenopus* embryo's has shown a similar but wider expression pattern for Fgf-3, suggesting a conserved, and perhaps more extensive, role in amphibian development [15].

Several embryonal stem and carcinoma cell lines demonstrate properties akin to cells of the early mouse conceptus [16, 17]. For example, F9 cells treated with retinoic acid (RA) and dibutyryl cAMP (dbcAMP) differentiate from an embryonal carcinoma stem cell to a parietal endoderm phenotype, and concomitantly induce Fgf-3 gene expression. These and similar cell types have provided *in vitro* systems to characterise the structure of Fgf-3 transcripts, which initiate from three promoter regions and terminate at either of two polyadenylation sites,

*To whom correspondence should be addressed

⁺Present address: Departament de Genetica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal, 08071 Barcelona, Spain

yielding a complexity of six classes of RNA species [11, 18]. In a previous report we showed that all mRNA species have a similar half life of between 120 and 150 minutes, suggesting that the choice of polyadenylation site does not significantly alter mRNA stability [19]. In addition we showed that a 1.7kb DNA fragment encompassing the three promoter regions was able to act as an inducible promoter for a marker gene when stably integrated into F9 cells. In the present paper we have used deletion analysis and DNase-I footprinting to identify both positive and negative regulatory elements which affect transcription from the Fgf-3 promoters.

MATERIALS AND METHODS

Cell culture

Mouse embryonal carcinoma cell lines F9 and PCC4 [20, 21] were cultured and induced to differentiate with RA and dbcAMP as described previously [18].

Extraction of cellular RNA and RNase protection analyses

Total RNA was isolated from pools of stable F9 transfectants and analysed by RNase protection as described [18]. Antisense ³²P-labelled RNA probes were synthesized from DNA plasmids using SP6 RNA polymerase under conditions recommended by the supplier (Promega). Probe 1 was prepared from plasmid CATR1-Fgf-3, which contained nucleotides 1–1688 fused to approximately 250 nucleotides of CAT coding region cloned in the vector Gem 4. For RNA synthesis the plasmid was linearized at nucleotide 1345 through a Sma-I site as shown in Figure 2a. Probe 2 contained nucleotides 231–1221 and was made from plasmid VC1013 as described [18, 22].

Plasmid constructs

Construction of plasmid pFgf-3/CAT has been described and was previously designated p1681CAT [19]. A preliminary series of unidirectional progressive 5' deletions spanning the 1.7 kb region was created using Exonuclease III under conditions described by the supplier (Gibco-BRL). The resultant single-stranded DNA was removed by treatment with SI nuclease, the ends blunted with Klenow enzyme, and the deleted DNAs recircularized.

A second set of progressive 5' deletion mutants were constructed using the polymerase chain reaction (PCR) with selected oligonucleotides based on the DNA sequences around regions of DNase-I protection. Sense oligonucleotides were 28 bases in length including 8 bases of a 5' BamHI site. The common antisense primer consisted of nucleotides 1681–1688 plus 12 bases of the vector including XhoI site to facilitate cloning. PCR products were ligated into BamHI/XhoI digested pFGF-3/CAT.

A series of directed deletion mutants was constructed by adding the missing upstream sequences also obtained as PCR fragments. Various 3' antisense oligonucleotide primers were designed to include a BamHI site to allow joining. The common 5' oligonucleotide used consisted of nucleotides 3 to 22 with a 5' HindIII site. The PCR fragments were inserted into the appropriate deletion mutant constructs. For deletion mutants $\Delta 705-719$, $\Delta 817-842$ and $\Delta 1058-1064$ the GATC of the BamHI site was removed by digestion with Mung bean Nuclease following BamHI digestion. This resulted in one or two extra nucleotides remaining at the deletion site.

CMV-LUC was constructed by inserting a HindIII/SacI fragment containing the firefly luciferase coding region from pJD201 [23], into pJ7 Ω [24].

All plasmid constructs were verified by DNA sequencing using Sequenase (USB).

Transfection of F9 cells

F9 cells for transfection were subcultured 1:5 (approximately 5×10^6 cells per 10cm dish) and medium changed 6 hours later. For transient transfections into differentiated F9 cells a calcium phosphate-DNA precipitate containing 15 μ g of supercoiled test plasmid DNA and 5 μ g of CMV-LUC DNA was placed on the cells and left overnight. The following day cells were washed and medium changed. Cells were harvested 24 hours later at which point they had received a total of 3 days treatment with RA and dbcAMP. Pools of undifferentiated stable transfectants were isolated as described [19].

CAT and luciferase assays

Pools of stable transfectants were subcultured at a ratio of 1:40 and either left untreated or exposed to retinoic acid and dbcAMP for 4 days as described above. Cell extracts were prepared as described previously [19]. CAT assays were carried out on extracts containing 100 or 200 μ g of protein as determined before heat treatment [25]. For transiently transfected cells 20 μ l of cell extract was used immediately for a luciferase assay. Luciferase activity was measured in 350 μ l of a buffer containing 25mM glycylglycine (pH7.8), 5mMATP, 15mM MgSO₄ using a luminometer (LKB 1251) following introduction of 33 μ l of 3mM luciferin solution [23]. The luciferase measurements were used to normalize transfection efficiency.

Preparation of DNA fragments for DNase I footprinting

Plasmids containing Fgf-3 genomic sequences were linearised with appropriate restriction enzymes, and treated with calf intestinal phosphatase (Boehringer-Mannheim). They were then labelled with γ -³²P-ATP (3000 Ci/mmol, Amersham) using T4 polynucleotide kinase (Pharmacia), and digested with a second restriction enzyme to generate various DNA fragments labelled at one end. For regions lacking appropriate enzyme sites, DNA probes were made by the polymerase chain reaction using the following primers; A (nts 601–625) and A' (nts 900–876); B (nts 801–825) and B' (nts 1100–1076); E (nts 1401–1425) and E' (nts 1700–1676). These primers were end labelled and used to make the DNA probes PCR-A', PCR-B, PCR-B' and PCR-E' respectively.

DNase I footprinting

DNase I footprinting assays [26] were performed as described by [27] with slight modifications in that binding reactions containing 1–10ng end-labelled DNA probes were allowed to proceed for 20 minutes at room temperature, followed by the addition of 3 U of DNase-I (Promega) and further incubation for 45 seconds. Reactions containing no nuclear extracts were digested with 1 U of DNase-I for 20 seconds. For DNA probes derived from restricted DNA, a 'G' and 'A+G' DNA sequence ladder generated as described by [28] from the same end-labelled DNA was used to identify the protected regions. For PCR derived probes, a DNA sequence ladder was generated using the end-labelled primer and Sequenase (USB).

Preparation of nuclear extracts

Nuclear extracts were prepared from undifferentiated F9 or PCC4 cell cultures, or after 4–5 days (F9) or 2 days (PCC4) treatment

with retinoic acid and dbcAMP, by the method described by [29] or according to a modified shorter version [27] with identical results.

RESULTS

The Fgf-3 promoter is responsive to RA and cAMP in F9 cells

We have previously shown that a fragment of DNA encompassing 1.7kb of sequence immediately 5' of the Fgf-3 coding region functions as a promoter [19]. Thus in F9 cells stably transfected with a construct incorporating this sequence, promoter activity was increased several fold after treatment with retinoic acid and dbcAMP. Using the same construct pFgf-3/CAT (Fig. 1b), we now show that the major response is to RA (5 fold) but that dbcAMP enhances both basal and RA induced responses (Fig. 1d). This mimics the increase of endogenous Fgf-3 transcripts seen in F9 cells after treatment with RA and/or dbcAMP [18, 19]. To determine whether sequences in the distal part of the 1.7kb Fgf-3 fragment contribute to promoter activity, upstream nucleotides (204 to 897) were joined to proximal sequences of the thymidine kinase promoter (Fig. 1c). Addition of these sequences resulted in a 5 fold increase in CAT activity following RA and dbcAMP treatment (Fig. 1d), confirming that within these upstream sequences there are elements which contribute to promoter function, and which show RA/dbcAMP sensitivity.

The Fgf-3 gene has three proximal promoter regions indicated as P1, P2 and P3 in figure 1a. Initiation of transcription at P1 results in a four exon messenger RNA in which the 5' exon joins the genomic sequences 6 bp upstream of the major P3 cap site. To determine whether the observed CAT activity reflects an increased level of correctly initiated transcripts from Fgf-3/CAT, total RNA isolated from a pool of stable F9 cell transfectants, before and after 3 days treatment with RA and dbcAMP, was assessed by RNase protection (Fig. 2). The amounts of transcript initiated at individual promoters was determined using ³²P-

labelled anti-sense RNA probes which encompass the various cap site regions. Probe 1 (Fig. 2a) was derived from pFgf-3/CAT and contains part of the CAT coding region in addition to Fgf-3 sequences. This allows transcripts initiated from the acquired Fgf-3/CAT hybrid gene to be discriminated from those derived from the endogenous Fgf-3. In figure 2b, comparison of lanes 1 and 2, or 3 and 4 shows the appearance of a protected fragment of approximately 344bp which corresponds to the increased level of transcripts initiating from the endogenous P2 promoter when F9 cells are induced to differentiate. Transcripts initiating from the endogenous P1 and P3 promoters were not detected as they only protect a very short region of probe 1 (Fig. 2a). However, transcripts initiating from P1 and P3 contained in pFgf-3/CAT were detected as protected fragments of approximately the predicted sizes, 281bp and 275bp respectively (Fig. 2b, lane 4). These fragments were not protected by RNA from the control vector transfected cells (lanes 1 and 2), nor by RNA from the uninduced cells containing pFgf-3/CAT (lane 3). A fragment of approximately 607bp was protected by probe 1 with RNA from induced F9 cells containing pFgf-3/CAT. This is the expected size of the fully protected probe and could result from transcripts initiated at the P2 promoter of pFgf-3/CAT, or from readthrough transcripts initiating upstream of the promoters. A second ³²P-labelled anti-sense probe which includes approximately 1kb of sequences upstream of P2 and entirely encompasses the first exon 1 (Fig. 2a) was used to investigate the usage of promoter P1, and monitor the presence of transcripts initiated upstream of the promoter region. This probe contains no CAT sequences and therefore does not distinguish between the endogenous and hybrid Fgf-3 transcripts, but detects the multiple cap sites of P1 as a discrete set of protected fragments (Fig. 2c). Readthrough or non-spliced transcripts should be revealed as additional larger sized fragments. Lanes 1 and 2, shows a set of protected fragments representing basal and induced level of endogenous P1 promoter activity respectively. Similarly, lanes 3 and 4 show protected

Figure 1. The Fgf-3 promoter regions and the response to retinoic acid and dbcAMP. (a) Schematic illustration of the Fgf-3 gene showing exons (boxes), and the promoter regions, P1, P2 and P3. The multiple cap sites for P1 and P2 are represented as vertical lines and the shaded region is the open reading frame. A1 and A2 are alternative polyadenylation sites. (b) Expression plasmid pFgf-3/CAT which contains nucleotides 1 to 1688 of Fgf-3 (thick dark line) encompassing the promoter regions and exons 1a, 1b and 1 (boxed). These sequences are inserted into the polylinker of pBLCAT3 [42]. (c) Expression plasmid pFgf-3c/CAT2 contains nucleotides 204 to 897 of Fgf-3 inserted into the polylinker of pBLCAT2. Down stream of the polylinker is the tk promoter (-105 to +51) which contains a 'TATA box' (d) Relative levels of CAT activity obtained with cell extracts from stable transfectants of F9 cells containing pOCAT, pFgf-3/CAT, pBLCAT2, or pFgf-3c/CAT2. Extracts were prepared from undifferentiated cells or cells treated with RA, dbcAMP or RA and dbcAMP as indicated.

Figure 2. Promoter and cap site usage of pFgf-3/CAT determined by RNase protection. (a) Diagram depicting pFgf-3/CAT (see Fig. 1 for details), the location of probes 1 and 2 (stippled boxes) and the predicted sizes of the protected fragments (thick lines and numbers) from pFgf-3/CAT and the endogenous Fgf-3. Autoradiographs of 6% acrylamide gels showing the protected RNA fragments obtained with probe 1 (b) and probe 2 (c) respectively. Total RNA (20 μ g) was analysed from undifferentiated (lanes 1 and 3) and differentiated (lanes 2 and 4) F9 cells transfected with vector pOCAT (lanes 1 and 2) or pFgf-3/CAT (lanes 3 and 4). The sizes of the protected fragments ($\pm$ 5 nucleotides) were calculated by comparison with a known sequencing ladder run in parallel tracks.

fragments derived from acquired as well as endogenous P1 promoter activity before and after induction. The characteristic pattern of multiple protected fragments from endogenous P1 while being intensified by addition of transcripts from pFgf-3/CAT remains qualitatively identical (compare lanes 2 and 4), indicative of correct cap site usage at the transfected P1 promoter, with no detectable readthrough transcripts. It was important to determine the absence of readthrough transcripts, since a protected fragment of 271bp (lanes 1 and 2, Fig. 2b) appeared to derive from fortuitous non-regulated transcription in control cells transfected with a CAT gene construct lacking a promoter. However, since such transcripts did not result in detectable CAT activity in these cells they appeared not to function as mRNAs. Nevertheless, a difference between transcription patterns for endogenous and acquired Fgf-3 promoters was noted. The endogenous Fgf-3 gene uses predominantly the P3 promoter in F9 cells [18, 19], while transcripts from the transgene initiate primarily at P1. Taken together these data show that the 1.7kb

DNA fragment endows regulated promoter activity with correct cap site usage, although additional sequences may be necessary for the balanced control of transcription from the three promoters.

5' Deletion analysis of the Fgf-3 promoter region

To identify more precisely regions of upstream DNA that regulate Fgf-3 gene expression, the plasmid pFgf-3/CAT was subject to a 5' deletion analysis starting from the BamHI site at nucleotide position 1 [22]. Deletion of the first 600 nucleotides had only a small effect on CAT activity following transient transfection of the mutant DNAs into differentiated F9 cells. However, deletion of the region between nucleotides 639 and 842 resulted in an 80% loss of CAT activity, suggesting the presence of a strong positive regulatory element between these nucleotides (data not shown). To provide a more accurate location of potential regulatory elements, another set of progressive deletion mutants of pFgf-3/CAT was generated using PCR primers as described in the methods section. The constructs with increasing degrees of deleted 5' sequence were transiently transfected into differentiated F9 cells, together with a CMV-Luciferase plasmid to obtain a correction for transfection efficiency. As shown in figure 3, deletion mutants which remove sequences up to nucleotide 460 have no significant effect on CAT activity compared to the parental pFgf-3/CAT. Deletion of sequences up to nucleotide 639 caused a modest drop in CAT activity, which with minor fluctuations was maintained for deletions up to nucleotide 703. However, deletion of another 17 bases to nucleotide 720, lead to a plasmid giving very poor CAT activity. Further deletions resulted in a relative decrease of activity between 769 and 843, and small increases in activity were noted comparing mutants 843 with 884 and 1044 with 1065, indicative of potential negative regulatory elements. Due to the low transfection efficiency of F9 cells, detection of regulatory elements between nucleotide 720 and the P3 promoter (nt 1677) were made difficult, because of relatively low CAT activities. We therefore used DNase-I footprinting analysis to locate candidate sequences in this region which could be subsequently tested for activity by targeted deletion analysis.

DNase-I footprinting

Nuclear extracts, prepared from undifferentiated and differentiated F9 and PCC4 cells were used to protect a series of DNA fragments spanning the presumed promoter domains as suggested by the preliminary 5' deletion analysis. PCC4 cells [20], are also mouse embryonal carcinoma cells which express Fgf-3 but they show faster induction kinetics and a preference for transcriptional initiation at promoter P2 [18]. The various DNA fragments to be used as substrates for footprinting were isolated as restriction enzyme sub-fragments or obtained by PCR amplification, and 32 P-labelled at the 5' end as described in the materials and methods.

Five regions of protection were observed within the 300bp immediately upstream of promoter P1. This region encompasses the sequences found in the preliminary deletion analysis to be responsible for a dramatic loss of promoter activity. The footprints, designated PS-1 to PS-5 (for Protected-Sequence 1 to 5), are shown in figure 4. Two regions, PS-1 and PS-2 were observed with extracts from both differentiated and undifferentiated F9 and PCC4 cells, although the protections were more marked and extensive when differentiated cell nuclear extracts were used (Fig. 4a). However, the PS-3 region was very evident with extracts from differentiated PCC4 cells, but only

Figure 3. 5' Deletion analysis of the Fgf-3 promoter. The left side of the figure shows a schematic diagram depicting the extent of a set of 5' deletion mutants and their relationship to the identified DNase-1 footprints. In the central region deletion mutants are designated by the nucleotides remaining in the plasmid. On the right side, are compiled results from several transient transfection experiments in differentiated F9 cells which show the relative activity of the constructs compared to the parent plasmid pFgf-3/CAT. In all experiments 5 μ g of a CMV-Luciferase plasmid was co-transfected to allow correction for transfection efficiency.

Figure 4. DNase I footprinting analysis of the region upstream of P1. (a) A SacI-AflII (595 to 897) fragment end-labelled at the AflII site which includes most of the region upstream of exon 1a was used as the target for footprinting. The DNase I cleavage patterns were generated either in the absence or presence of nuclear extracts from differentiated or undifferentiated, F9 or PCC4 cells. The regions of protected sequence PS-1, PS-2 and PS-3 are indicated. (b) and (c) Probes PCR-A' and PCR-B spanning bases 601–900 and 801–1,100 respectively, were labelled either at nt 900 (PCR-A') or nt 801 (PCR-B). Nuclear extracts are as described above, but the protected regions were determined by comparison with sequencing tracks obtained using the same labelled primer as described. The regions corresponding to PS-4 A to PS-5 are indicated.

poorly visible with F9 cell extracts, suggesting that it may be associated with activity from P2 which predominates in these cells [18, 19]. An extensive region of protection was revealed from nucleotide 817 to 898. This was divided into three sub-regions

designated PS-4A, PS-4B and PS-5 (Fig. 4b and c). For each of these regions there was no clear difference in the extent of protection using nuclear extracts from differentiated or undifferentiated cells.

Figure 5. DNase I footprinting analyses of the region between promoter P1 and P3. Probes PCR-B' and PCR-E' are PCR amplified fragments spanning bases 801–1100 and 1401–1700, and labelled at nt 1100 (PCR-B') and nt 1700 (PCR-E') respectively. Nuclear extracts and markers are as described in figures 4 and 5. Protected regions corresponding to PS-6 to PS-9 (a) and PS-12 and PS-13 (b) are indicated.

A further six footprints were identified in the region between the promoters P1 and P2 (nucleotides 900–1300), and were designated PS-6 to PS-11. Of these, PS-6 to PS-9 were detected with extracts from both cell types (Fig. 5a), whereas PS-10 and PS-11 were only observed in PCC4 cells (data not shown). In addition, there were several regions where the intensities of specific bands differed somewhat from those in the control DNase-I ladder, and could be indicative of other binding factors. In the absence of additional information these have not been given specific designations. In the region between promoters P2 and P3, (nucleotides 1300–1600), two potential binding sites, PS-12 and PS-13 were identified with extracts from differentiated cells (Fig. 5b). These regions were better protected with extracts from differentiated F9 cells, compared with differentiated PCC4 cells (data not shown). In all, at least thirteen regions of strong protection were found within, and proximal to, the Fgf-3 promoter domains. A summary of the protected regions are marked on a sequence of the region (nucleotides 600 to 1700) together with the position of the exons (Fig. 6).

Targeted deletion analysis

Based on the 5' deletion and DNase-I footprinting results, mutant plasmids were constructed in which only small regions of

Figure 6. Scheme depicting the positions of PS-1 to PS-13 in relation to the Fgf-3 promoter regions. The Fgf-3 genomic sequence numbers and the position of exons (stippled boxes) are as previously described. The protected sequences PS-1 to PS-13 are boxed and labelled, arrows depict the major cap sites and the splice donor and acceptor site.

promoter sequence were removed (Fig. 7). Constructs were transiently transfected into differentiated F9 cells with the control CMV-Luciferase plasmid, and results standardized for luciferase activity were expressed relative to pFgf-3/CAT. In agreement with the previous results, removal of sequences encompassing the footprinted region PS-1 ($\Delta 658-702$) had little effect on the resultant CAT activity. In contrast, removal of sequences which included the footprinted region PS-2 ($\Delta 705-719$), or sequences between 817 and 855 ($\Delta 817-842$ and $\Delta 829-855$) which include the protected sequence PS-4A, resulted in transfected F9 cells with dramatically reduced levels of CAT activity. Several other deletion mutants also resulted in a comparatively lower level of CAT activity. These included plasmids with loss of sequences between 881–898, 1016–1043 and 1530–1543 which encompassed the footprinted regions designated, PS-5, PS-7/8 and PS-12 respectively. However, for these three deletion mutants the reduction in CAT activity in transiently transfected F9 cells was modest. Interestingly, the loss of sequences between 857 and 866 encompassing a footprinted region designated PS-4B which lies between PS-4A and PS-5, resulted in a 3.5 fold increase in CAT activity. This suggests the presence of a negative regulatory element which appears to suppress promoter activity in both undifferentiated and differentiated F9 cells. Moreover, deletion of sequences corresponding to footprints PS-9 and PS-13 also resulted in a stimulation of CAT activity of 2 fold and 3.7 fold respectively, indicative of other negative regulatory elements. However, not all the footprints could be accounted for in the functional analysis, since some clearly defined regions of DNase-I protection showed no significant effect on CAT activity when deleted. In particular these included mutants lacking regions PS-6, PS-10 and PS-11. Nevertheless, taken together these findings suggest the presence of at least two positive and three negative regulatory elements that affect the Fgf-3 promoters in F9 cells.

The deletion analyses described above relied on the activity of the CAT reporter plasmid following transient transfection of various constructs into differentiated F9 cells. These assays cannot completely reflect the regulation of the Fgf-3 gene since expression is not sensitive to RA and dbcAMP induction. However, F9 cells which contain stable integrated copies of the parental plasmid pFgf-3/CAT show an increase in promoter activity following RA and dbcAMP induced differentiation (Fig.

Figure 7. Deletion analysis of specific footprinted regions. The left section shows a scheme of the deletions which are marked by a gap in the line depicting each plasmid. The mutants are listed in the center designated by the missing nucleotides. On the right the results of the corrected transient transfection in differentiated F9 cells are shown as a histogram compiled from several experiments and normalized to the activity of Fgf-3/CAT.

1), as does the endogenous Fgf-3 gene. Such cell cultures also showed much higher overall levels of CAT activity than the transient transfectants where only a small proportion of cells in the culture acquire the test plasmid. Therefore we assessed the effects of specific deletions on the activity of Fgf-3 promoters in both undifferentiated and differentiated F9 cells containing integrated copies of the mutant plasmids. Routinely, transfected cultures were selected to be comprised of more than a 100 independent cell colonies. This was to minimise effects caused by position of integration and/or variations in copy number for individual colonies of the pool. The similarity in average copy number of the cell populations was confirmed using Southern blot analysis on doubling dilutions of total cellular DNA (data not shown). CAT activity obtained for each pool of cells was corrected for total protein content and expressed as units of CAT activity per 100 μ g protein (Fig. 8). As expected CAT activity was substantially higher in the stable transfected cell pools, compared with equivalent transient transfectants. Nevertheless, the results obtained using transient transfections were substantiated, and extended to show the effect of RA and dbcAMP induced differentiation. In comparison with pFgf-3/CAT, mutants lacking sequences for footprinted regions PS-4A (Δ 817–842, Δ 829–855) and PS-5 (Δ 881–898) showed very low basal and induced levels of CAT activity indicative of potent positive elements (Fig 8a). In contrast those lacking PS-4B (Δ 857–866), PS-9 (Δ 1058–1064) and PS-13 (Δ 1562–1570) showed much greater activity than pFgf-3/CAT in both cell states, consistent with the presence of strong negative regulatory elements (Fig 8b). The elevation of both basal and induced levels of transcription suggests that the effect of these negative elements is manifest in differentiated and undifferentiated F9 cells. Apart from deletions encompassing PS-4A, none of the mutations caused loss of inducibility arising from differentiation with RA and dbcAMP, although the degree of induction (ratio of induced to uninduced activity) was changed to a small extent in some cases. A comparison of the results from the transient and stable transfectants revealed one contradictory finding (Figs 7 and 8). The deletion mutant missing sequences of the footprinted region PS-2 (Δ 705–719), showed a dramatic reduction in CAT activity when compared to pFgf-3/CAT in a transient transfection assay, whereas the stable transfectants containing these constructs gave comparable results.

DISCUSSION

In situ hybridization studies have shown that during embryogenesis, Fgf-3 transcripts are expressed at multiple sites and in a defined temporal sequence [10, 12, 13], leading to the idea that Fgf-3 may have multiple roles in mammalian development. The strict spacio-temporal pattern of expression is likely to be achieved in large part by regulation of promoter activities. To understand this regulation, we have begun to analyse the Fgf-3 promoter in order to identify the regulatory elements which control gene activity. Previous studies to determine the structure of the gene transcripts used embryonal carcinoma and stem cell lines to provide a suitable *in vitro* expression system [11, 18]. Few other cell lines, or adult tissues, have been found that transcribe Fgf-3, with the exception of some mouse breast carcinoma cells in which transcription has been induced by the presence of a proviral element. The results from all these analyses, identified at least three distinct sets of cap sites, suggestive of three separate but not necessarily independent promoters (see Fig. 1a). Previous transfection studies showed that F9 embryonal carcinoma cells were the only cell line tested able to support expression from Fgf-3 promoter constructs [19]. Following transient transfection both differentiated and undifferentiated F9 cells supported similar Fgf-3 promoter activities. This was in contrast to stable transfectants where expression was markedly induced upon differentiation, which parallels the activity observed for the endogenous gene. Induction of F9 cell differentiation with RA and dbcAMP results in a change from embryonal stem cell to parietal endoderm, which is the first site of Fgf-3 expression detected in the conceptus [12, 21, 30]. We have shown that RA is more critical than cAMP for Fgf-3 promoter activity (Fig. 1d). However, examination of upstream sequences has not revealed a consensus binding site for retinoic acid receptor (RAR). Both for the endogenous gene and artificial Fgf-3 constructs, induction of transcripts is not immediate but occurs 24 to 48 hours after treatment with RA ([18] and data not shown), suggesting its indirect effect on gene expression, possibly through transcriptional regulation of other transcription factors. Consistent with this idea are previous reports that treatment of F9 cells with RA alters the binding properties of several transcription factors (reviewed in [31]). For example, binding activity for RA β -receptor, AP1, TCIIA and TCIIIB are reportedly

Figure 8. Effect of specific deletions on the Fgf-3 promoter activity in differentiated and undifferentiated F9 cells. (a) The same deletion mutants described in figure 7 were transfected into F9 cells and stable cell lines selected as described. Cells were grown in the presence or absence of retinoic acid and dbcAMP for 72 hours and the CAT activity measured in extracts normalized to 100 mg of protein. (b) Deletion mutants with CAT activities off-scale in (a) plotted on a smaller scale. In both (a) and (b) open and filled columns represent undifferentiated and differentiated F9 cell cultures respectively.

elevated, while PEA2, the ATF family and E2F are reduced.

To locate potential transcription factor binding sites we have identified, by means of deletion analysis, regions of sequence that are functionally important for expression of the reporter gene CAT under the control of Fgf-3 upstream sequences. The locations of these regions were further defined by DNase-I footprinting and subsequent site directed deletion analysis. In a 1kb span encompassing the three promoter regions, at least 14 potential binding sites for regulatory factors were identified. Of these, nine DNase-I protected sites were shown to have some effect on expression from the Fgf-3 promoter in F9 cells. The sequence immediately upstream of P1 (nts 817–898) contained an extended region of protection which encompassed PS-4A, PS-4B and PS-5. Deletion of the PS-4A footprint virtually destroyed all promoter activity in both transient and stable transfectants, indicating that it is vital for at least P1 and P3 activity. The sequence (PS-4A) shows similarity but not consensus identity, to other transcription factor binding sites such as AP1, and a consensus identified for RA mediated (RA receptor independent) activation of c-jun in F9 cells [32, 33]. The

transcription factor c-jun binds as a homodimer or heterodimer with members of the fos family to AP1 sites and begins to increase in F9 cells within 12–24 hours of exposure to RA, slightly earlier than Fgf-3 [32]. However, although the kinetics of c-jun induction would be consistent with Fgf-3 promoter activation, the footprint (PS-4A) was found with extracts from both differentiated and undifferentiated cells, suggesting protection was not mediated by an induced factor. Nevertheless, exchange of protein binding factors giving similar footprints cannot be excluded. Alternatively, the bound protein may be modified or associate with an induced accessory protein to generate a complex which stimulates promoter activity.

The footprint PS-4B is juxtaposed to PS-4A and confers a strong repression on transcription, since its deletion results in a substantial increase in CAT activity in both differentiated and undifferentiated F9 cells. Furthermore, there was no detectable difference in the footprint pattern using extracts from differentiated and undifferentiated F9 cells. We therefore presume that RA induction of promoter activity results from an over-riding positive effect, rather than relief of the negative effect mediated through this regulatory element. In transiently transfected F9 cells the stimulation of CAT activity resulting from deletion of PS4B was quite modest (3.5 fold), while for the stable transfectants the increase was more than 19 fold (compare Figs. 7 and 8). This difference may in part be due to limiting amounts of factors in the few cells which take-up DNA in transient transfections. Consequently, these transfected cells are unable to generate a full response in the absence of the negative regulatory element. Such a situation is unlikely to arise in stable transfectants which contain fewer copies of the transfected gene. The PS-4B sequence, shows some similarity to PS-9 which also has silencer activity, but at a much reduced level. Both elements are rich in cytidine nucleotides, but unlike PS-4B, PS-9 contains a consensus sequence for the transcription factor AP2 [34]. The expression of AP2 is elevated following RA stimulation of the human teratocarcinoma cell line NT2, and similar activity might be expected for F9 cells [35]. However, involvement of AP2 in regulation of Fgf-3 promoter activity seems unlikely because the PS-9 footprint is similarly intense with extracts from both differentiated and undifferentiated cells. Furthermore, AP2 is generally recognised as a positive regulatory factor whereas PS-9 sequences are associated with moderate silencer activity. Thus AP2 is a rather weak candidate for interaction with PS-9 and we suspect other factor(s) may be involved.

A third footprinting region demonstrating silencer activity, PS-13, shows no sequence relationship to PS-4B or PS-9, but fits the consensus sequence for the GATA family of transcription factors. Surprisingly, the protection is much stronger with extracts from differentiated than undifferentiated cells, which is not consistent with direct binding of a negative regulatory protein. At least four mammalian members of the GATA family of transcription factors have been described, and one GATA-4 is induced in F9 cells following RA induced differentiation (see [36]). In the majority of cases, the GATA factors act as positive regulators of transcription. However, a mutation in a GATA-1 binding site in the mouse γ -globin promoter prevents the extinction of this gene in adult life. The mechanism of this transcriptional extinction is obscure, but GATA-1 may function by excluding the binding of a potent activator, or alternatively, diminish the potential interaction between the γ -globin gene and the locus control region [37, 38]. A possible explanation to account for the data presented here is that the weak footprint

found in undifferentiated cells reflects binding of a negative regulatory factor which is replaced or modified resulting in a less potent effect in differentiated cells. However, as with the effect of deleting PS-4B, a dramatic increase in promoter activity is seen in both undifferentiated and differentiated cells suggesting the element is operative in both cell states. Perhaps it is the association of the factor with accessory proteins that modulates its activity, which is reflected in the more distinct footprint seen with differentiated cell extracts. Interestingly, footprint PS-12 also contains the GATA consensus, but the footprint is weak with extracts from both differentiated and undifferentiated cells, and deletion of this regions has little effect on CAT activity.

In transient transfection assays deletion of PS-2 results in a substantial decrease in CAT activity, which was not observed in comparable stable transfectants. At present we have no specific explanation for this observation, but suspect the high concentration of input DNA in transient transfections must be distorting the balance of transcription factors in an ill-defined manner. It is interesting to note that the footprint encompasses a core ets-1 binding site adjacent to a CBF site, a combination previously described for the T-cell receptor β gene [39]. Analogous partnerships between ets-1 and AP1 have also been described for other genes [40, 41]. Although it is unclear whether this site is important for transcriptional regulation in F9 cells, the multiplicity of embryonic expression sites for Fgf-3 might argue that the importance of the PS-2 region is context dependent. This speculation could be extended to the other sites which show footprinted regions with F9 and/or PCC4 nuclear extracts but demonstrate little or no significant effect when deletions are tested in F9 cells. Future work will be directed at examining the regulation of this gene in other cell types, as well as identifying the protein factors that bind to the elements defined here. To provide further insights into the processes of tissue specific regulation in the embryo we would like to extend this cell culture analysis to transgenic mice, so as to monitor the effects of these elements on the spatial and temporal expression of Fgf-3 in the developing embryo.

ACKNOWLEDGEMENTS

We thank Steven Goodbourn and Nicholas Jones for comments on the manuscript, and Anne-Marie Florence for additional technical help.

REFERENCES

- Burgess, W. and T. Maciag. (1989). *Annu. Rev. Biochem.* 58, 575–606.
- Dickson, C., R. Deed, M. Dixon and G. Peters. (1989). *Prog. Growth Factor Res.* 1, 123–132.
- Baird, A. and M. Klagsbrun. (1991). *Ann N Y Acad Sci.* 638, xiii–xv.
- Tanaka, A., K. Miyamoto, N. Minamino, M. Takeda, B. Sato, H. Matsuo and K. Matsumoto. (1992). *Proc. Natl. Acad. Sci. USA.* 89, 8928–8932.
- Dickson, C., R. Smith, S. Brookes and G. Peters. (1984). *Cell.* 37, 529–536.
- Peters, G., S. Brookes, R. Smith and C. Dickson. (1983). *Cell.* 33, 369–377.
- Muller, W., F. Lee, C. Dickson, G. Peters, P. Pattengale and P. Leder. (1990). *EMBO J.* 9, 907–913.
- Ornitz, D., R. Cardiff, A. Kuo and P. Leder. (1992). *J. Natl. Cancer Inst.* 84, 887–892.
- Stamp, G., V. Fantl, R. Poulson, S. Jamieson, R. Smith, G. Peters and C. Dickson. (1992). *Cell Growth Differ.* 3, 929–938.
- Jakobovits, A., G. Shackleford, H. Varmus and G. Martin. (1986). *Proc. Natl. Acad. Sci. USA.* 83, 7806–7810.
- Mansour, S. and G. Martin. (1988). *EMBO J.* 7, 2035–2041.
- Wilkinson, D., G. Peters, C. Dickson and A. McMahon. (1988). *EMBO J.* 7, 691–695.
- Wilkinson, D., S. Bhatt and A. McMahon. (1989). *Development.* 105, 131–136.
- Mansour, S. L., J. M. Goddard and M. R. Capecchi. (1993). *Development.* 117, 13–28.
- Tannahill, D., H. V. Isaacs, M. J. Close, G. Peters and J. Slack. (1992). *Development.* 115, 695–702.
- Hogan, B., D. Barlow and R. Tilly. (1983). *Cancer Surveys.* 2, 115–140.
- Darrow, A., R. Rickles and S. Strickland. (1990). *Methods in Enzymology.* 190, 110–140.
- Smith, R., G. Peters and C. Dickson. (1988). *EMBO J.* 7, 1013–1022.
- Grinberg, D., J. Thurlow, R. Watson, R. Smith, G. Peters and C. Dickson. (1991). *Cell Growth and Diff.* 2, 137–143.
- Jakob, H., T. Boon, J. Gaillard, J.-F. Nicolas and F. Jacob. (1973). *Ann. Microbiol. (Institut Pasteur).* 124B, 269–282.
- Strickland, S., K. Smith and K. Marotti. (1980). *Cell.* 21, 347–355.
- Moore, R., G. Casey, S. Brookes, M. Dixon, G. Peters and C. Dickson. (1986). *EMBO J.* 5, 919–924.
- de Wet, J., K. Wood, M. DeLuca, D. Helinski and S. Subramani. (1987). *Mol. Cell. Biol.* 7, 725–737.
- Morgenstern, J. and H. Land. (1990). *Nucleic Acids Res.* 18, 3587–3596.
- Sleigh, M. (1986). *Anal. Biochem.* 156, 251–256.
- Galas, D. and A. Schmitz. (1978). *Nucleic Acids Res.* 5, 3157–3170.
- Lee, K., A. Bindereif and M. Green. (1988). *Gene Anal. Techn.* 5, 22–31.
- Maxam, A. and W. Gilbert. (1980). *Meth. Enzymol.* 65, 497–559.
- Dignam, J., R. Lebovitz and R. Roeder. (1983). *Nucleic Acids Res.* 11, 1475–1489.
- Strickland, S. and V. Mahdavi. (1978). *Cell.* 15, 393–403.
- Sleigh, M. (1992). *BioEssays.* 14, 769–775.
- Kitabayashi, I., Z. Kawakami, R. Chiu, K. Ozawa, T. Matsuo, S. Toyoshima, K. Umehara, R. Evans, G. Gachelin and K. Yokoyama. (1992). *EMBO J.* 11, 167–175.
- Risse, G., K. Jooss, M. Neuberger, H.-J. Brüller and R. Müller. (1989). *EMBO J.* 8, 3825–3832.
- Williams, T. and R. Tijian. (1991). *Genes & Develop.* 5, 670–682.
- Luscher, B., P. Mitchell, T. Williams and R. Tijian. (1989). *Genes & Develop.* 3, 1507–1517.
- Arceci, R., A. King, M. C. Simon, S. Orkin and D. Wilson. (1993). *Mol. Cell. Biol.* 13, 2235–2246.
- Berry, M., F. Grosfeld and N. Dillon. (1992). *Nature.* 358, 499–502.
- Crossley, M. and S. Orkin. (1993). *Curr. Opin. in Gen and Develop.* 3, 232–237.
- Prosser, H., D. Wotton, A. Geggion, J. Ghysdael, S. Wang, N. Speck and M. Owen. (1992). *Proc. Natl. Acad. Sci. USA.* 89, 9934–9938.
- Gutman, A. and C. Wasylyk. (1991). *Trends in Genetics.* 7, 49–54.
- Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. Leprince and D. Stehelin. (1990). *Nature.* 346, 191–193.
- Luckow, B. and G. Schütz. (1987). *Nucleic Acids Res.* 15, 5490.