Identification of a key regulatory element for the basal activity of the human insulin-like growth factor II gene promoter P3

Luc E. G. RIETVELD¹, P. Elly HOLTHUIZEN and John S. SUSSENBACH

Laboratory for Physiological Chemistry, Utrecht University, Stratenum, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands

Transcription of the human insulin-like growth factor II (IGF-II) gene is under the control of four promoters (P1–P4) that are differentially active during growth and development. Promoter 3 (P3) is the most active promoter during fetal development as well as in most adult tissues. P3 is also the most active promoter in tumour tissues and cell lines expressing IGF-II. Transient transfections of HeLa and Hep3B cells with truncated promoter constructs revealed that the region between –289 and –183 relative to the transcription start site supports basal promoter activity in both cell lines. Footprint experiments showed that the region between positions –192 and –172 (P3-4) is the only element bound by nuclear proteins. P3-4 is bound by five proteins, of which three proteins (proteins 3, 4 and 5) bind specifically and are expressed at the same levels in HeLa and Hep3B cells.

Electrophoretic mobility shift assays and differential footprint experiments revealed the presence of two protein-binding regions within the P3-4 element. Proteins 4 and 5 bind box A (–193 to –188), whereas box B (–183 to –172) is bound by protein 3. From transcription experiments *in vitro* it can be concluded that Box A is essential for P3 activity. Box A is part of a region 11 dG residues long and is protected by proteins 4 and 5 that bind a contiguous set of six dG residues. DNA-binding of proteins 4 and 5 to box A requires the presence of Zn²⁺ ions. Thus structural and functional analysis reveals that the P3-4 element is a key regulatory element of P3 that contains two separate binding sites for proteins essential for the basal activity of IGF-II P3.

INTRODUCTION

Human insulin-like growth factor II (IGF-II) is a polypeptide growth factor of 67 amino acid residues that plays an important role in fetal and post-natal development. It is expressed in various tissues but the main source is the liver [1]. The function of IGF-II *in vivo* became clear after targeted disruption of the gene in mice. Mice carrying IGF-II-null mutations exhibit a body weight decrease to 60 %, revealing the importance of IGF-II for fetal development [2]. IGF-II is also a potent stimulator of cell differentiation, as was shown for neuronal cells [3] and muscle cells [4]. In addition to its differentiation potential, IGF-II also stimulates the proliferation of many different cells in culture [5].

The structure of the IGF-II gene reveals that the gene is a complex transcription unit. The gene consists of nine exons, of which exons 1–6 are non-coding leader exons. Exons 7, 8 and the first part of exon 9 encode prepro-IGF-II, which consists of 180 residues. Exons 1, 4, 5 and 6 are preceded by distinct promoters (P1–P4), which give rise to a family of mRNA transcripts containing different leader exons but the same coding exons [6–8].

Transcription of the IGF-II gene is controlled in a tissue- and development-specific way. During fetal development promoter 3 (P3) is the most active promoter in many fetal tissues including fetal liver, whereas promoters 2 and 4 (P2 and P4) are moderately active. In adults these three promoters remain active in non-hepatic tissues, but in the liver they are down-regulated and another promoter, P1, is activated. In cell lines and tumour tissues expressing IGF-II, P3 is the most active promoter, giving rise to an mRNA of 6.0 or 2.2 kb depending on the poly(A) signal used.

Promoter P3 spans 1300 bp upstream of the transcription start site of exon 5 and is very G/C-rich, consisting of 70–80 % GC base pairs. Transient transfections in HeLa and Hep3B cells with truncated promoter constructs revealed that the P3 region that confers cell type-specific activity in Hep3B cells is located between positions –1300 and –289 [9]. From these transfection experiments it also became clear that the region downstream of position –289 supports basal promoter activity in both cell lines [9]. This proximal P3 region contains a TATA box and a CCAAT box. Several G/C box-binding transcription factors have been identified that are involved in P3 regulation via this proximal region. The product of the Wilms' tumour suppressor gene, WT1, is able to repress P3 by binding to different sites in the promoter and within exon 5 [10]. Another factor that binds the same site as WT1, EGR-1, is able to stimulate P3 [11]. In addition to the TATA box, we have identified four other regions in the proximal region that were bound by nuclear factors from HeLa or Hep3B cells in footprinting experiments. These regions were designated PE3-1, PE3-2, PE3-3 and PE3-4 (here called P3-1, P3-2, P3-3 and P3-4 respectively), located in the region between the transcription start site (+1) and position –195 [9,12]. The relative importance of these four proximal elements for transcription was established by transcriptions *in vitro* and transfections of HeLa and Hep3B cells employing truncated promoter constructs [9,12]. In both cell lines P3-1, P3-2 and P3-4 are essential for P3 activity, whereas P3-3 seems to be of minor importance. By employing DNase I footprinting analysis and electrophoretic mobility shift assay (EMSA) it has been shown that P3-1 (–89 to –64), containing an inverted CCAAT box, is bound by transcription factors of the family of CCAAT box-binding factors such as NF-1. P3-2 (–126 to –102) consists of a consensus G/C box that is bound by the ubiquitous tran-

Abbreviations used: AdML, adenovirus major late; ds, double-stranded; EMSA, electrophoretic mobility shift assay; IGF, insulin-like growth factor; MCS, multiple cloning site; WCE, whole cell extract.

¹ To whom correspondence should be addressed.

scription factor Sp1. Element P3-4, located around position -180, is essential for transcriptional activity of P3. P3-4 does not contain an obvious consensus sequence for any known transcription factor, suggesting that an as yet unidentified transcription factor binds to this element.

In the present paper we describe the characteristics of proteins that bind to the P3-4 element located in the region between -192 and -172 relative to the transcription start site. In addition, their importance for P3 activity has been investigated.

MATERIALS AND METHODS

Plasmids

The IGF-II promoter P3 expression constructs pFlag-Sma (positions -183 to -7) and pFlag-Xba (positions -289 to -7) were constructed as described previously [12]. Flag constructs contain both (1) a G-less cassette to monitor transcription rates *in vitro* and (2) a luciferase reporter gene to monitor promoter activity by transient transfection assays. The pAdML200 clone containing the adenovirus major late (AdML) promoter was used as an internal control template for transcriptions *in vitro* and was derived from pMLC₂AT19 [13], which contains adenovirus major late promoter sequences -400 to +10 linked to a 200 bp G-free cassette. pTK-Luc was made by the insertion of the -105 to +51 region of the Herpes Simplex Virus Thymidine Kinase promoter linked to the luciferase gene in pSla3 [9]. Subsequently, pP3-4TK-Luc was constructed by cloning the double-stranded (ds) phosphorylated oligonucleotide P3-4TOT into the *Sma*I site of pTK-Luc.

Oligonucleotides

Oligonucleotides were synthesized with a Pharmacia/LKB Gene Assembler Plus and purified with a Sephadex G-50 column; ds oligonucleotides used in EMSA experiments and in transcriptions *in vitro* were purified on a non-denaturing 5% (w/v) polyacrylamide gel. Oligonucleotides that were used are as follows (mutations shown in italics): P3-4TOT upper, 5'-TCGAG-CCCCCCCCCGGGGCCCCAGGGCCCCCAGCC-3'; P3-4TOT lower, 5'-TCGAGGCTGGGGGCCCTGGGGCCCCCGGGGGGGGGC-3'; P3-4L upper, 5'-TCGAGCCCCCCCCCCCCGGGG-3'; P3-4L lower, 5'-TCGACCCCGGGGGGGGGGGC-3'; P3-4M upper, 5'-TCGAGCCCCGGGGCCCCAGGGC-3'; P3-4M lower, 5'-TCGAGCCCTGGGGCCCCCGGGC-3'; P3-4R upper, 5'-TCGACCCAGGGCCCCCAGCC-3'; P3-4R lower, 5'-TCGAGGCTGGGGGCCCTGGGG-3'; P3-4TOTM1 upper, 5'-TCGAGCCCTAATCCCCGGGGCC-CCAGGGCCCCCAGCC-3'; P3-4TOTM1 lower, 5'-TCGAGGCTGGGGGCCCTGGGGCCCCGGGATTAGGGC-3'; P3-4TOTM2 upper, 5'-TCGAGCCCCCCCCATTAGGCC-CCAGGGCCCCCAGCC-3'; P3-4TOTM2 lower, 5'-TCGAGGCTGGGGGCCCTGGGGCCTAATGGGGGGGGGC-3'; P3-4TOTM3 upper, 5'-TCGAGCCCCCCCCCGGGATTACAGGGCCCCCAGCC-3'; P3-4TOTM3 lower, 5'-TCGAGGCTGGGGGCCCTGTAATCCCGGGGGGGGGGC-3'; P3-4TOTM4 upper, 5'-TCGAGCCCCCCCCCGGGGCCCTAATCCCCCAGCC-3'; P3-4TOTM4 lower, 5'-TCGAGGCTGGGGGATTAGGGGCCCGGGGGGGGGGC-3'; multiple cloning site (MCS) upper, 5'-AGCTTCAGCTGCTAGCATCGATCGGCCGACATCTC-3'; MCS lower, 5'-CATGGAGATGTGCGCCGATCGATGCTAGCAGCTGA-3'; P3-1 upper, 5'-GATCCCCATTGGCGCGGGCGCGAGGA-3'; P3-1 lower, 5'-GATCTCTCGCGCCCGCGCCAATGGG-3'.

Cell culture and transient transfections

Hep3B cells [14] were cultured in α -modified minimal essential medium; HeLa cells were cultured in Dulbecco's modified Eagle's medium. Media were supplemented with 10% (v/v) fetal calf serum, 100 i.u./ml penicillin, 100 μ g/ml streptomycin and 300 μ g/ml glutamine. Introduction of expression constructs into Hep3B and HeLa cells was performed by means of a modified version of the calcium phosphate co-precipitation method [15]. HeLa and Hep3B cells were seeded at a density of 3×10^5 cells per 3-cm diam. dish. Fresh medium was added 2 h before transfection. Each dish received a total of 3 μ g of DNA consisting of 2 μ g of P3-reporter plasmid, 0.25 μ g of RSV-LacZ as an internal control and 0.75 μ g of pBluescript KS as carrier DNA. Cells were exposed to the DNA-calcium co-precipitate for 16 h. Fresh medium was added and after 24 h the cells were harvested for the determination of luciferase and β -galactosidase activities. Preparation of cell extracts for the β -galactosidase and luciferase assays were performed as described [15,16].

EMSA

Ds oligonucleotides were end-labelled by phosphorylation with [γ -³²P]ATP and T4 polynucleotide kinase [15]. EMSA binding conditions were determined as described [17]. Briefly, binding reactions were performed by mixing DNA probes, ds competitor DNA and binding buffer [20 mM Hepes/KOH (pH 7.6)/1 mM MgCl₂/75 mM KCl/1 μ g of poly(dI-dC) (Pharmacia, Uppsala, Sweden)/1 mM dithiothreitol/0.018% Nonidet P40], after which protein extract was added. The total reaction volume was 15 μ l; the mixture was incubated for 60 min on ice. Bound and unbound probe were separated on a 5% (w/v) polyacrylamide gel (ratio of acrylamide to bis, 37.5:1) in 0.5 $\times$ Tris/borate/EDTA. After the gel had been dried, the bands were detected by autoradiography.

DNase I footprinting

DNase I footprinting reactions were performed as described previously [17]. The probe for DNase I footprinting contained an *Eco*RI/*Sal*I P3 fragment from -290 to +134 labelled in the upper strand.

Preparation of cell extracts

For the preparation of whole cell extract (WCE), HeLa cells were grown in suspension cultures to a density of 5×10^5 cells/ml. Extract preparation was performed by the method of Manley et al. [18]. Partial purification of HeLa WCE was performed with the use of a phosphocellulose column (Whatman) as described previously [19]. Proteins were adsorbed on phosphocellulose and eluted with a continuous KCl gradient (0.04–1 M). Separate fractions were dialysed against buffer A100 [20 mM Hepes/KOH (pH 7.9)/100 mM KCl/20% (v/v) glycerol/0.5 mM EDTA/0.5 mM PMSF/1 mM dithiothreitol] and tested for protein binding. The protein concentrations were determined by the Bradford protein quantification assay (Bio-Rad, Brussels, Belgium). For the preparation of nuclear extracts HeLa cells were grown in suspension culture. Briefly, nuclei were isolated and washed three times with 25 mM Tris/HCl (pH 7.5)/10% (w/v) sucrose/1 mM PMSF/1 mM dithiothreitol. The nuclei were suspended in the same buffer also containing 1 mM EDTA. Nuclear proteins were extracted by the addition of 5 M KCl to a final concentration of 0.3 M. After centrifugation the final supernatant was adjusted to 0.1 M KCl by the addition of buffer A100 [20,21].

Transcription reactions *in vitro*

For transcription reactions *in vitro* plasmid DNA species were purified twice on CsCl gradients and were then dialysed extensively against 10 mM Tris/HCl (pH 8.0)/1 mM EDTA. A typical transcription reaction *in vitro* contained 75 ng of P3 template and 200 ng of the adenovirus major late promoter template (pAdML200) as an internal control. The amount of WCE used in each reaction was 125 μ g. Reactions were assembled on ice; for competition assays WCE was preincubated for 30 min on ice with a fixed amount of ds competitor oligonucleotide as indicated. [α - 32 P]CTP (5 μ Ci; specific radioactivity 400 Ci/mmol; Amersham, Little Chalfont, Bucks., U.K.), 0.06 mM ATP, 0.06 mM UTP, 0.01 mM CTP, 0.06 mM 3'-O-methyl-GTP, 16 units of RNasin (Promega, Madison, WI, U.S.A.), 5 mM MgCl₂ and 5 mM creatine phosphate (Boehringer, Mannheim, Germany) were added to each reaction. The total reaction volume was 20 μ l; the final KCl concentration was 60 mM. Transcriptions were performed for 50 min at 30 °C. After this, 10 units of RNase T1 (Boehringer) was added and incubated for an additional 10 min at 30 °C. Subsequently, 180 μ l of stop solution [20 mM Tris/HCl (pH 8.0)/0.5% SDS/100 mM LiCl/1 mM EDTA/7 M urea/100 μ g/ml *Escherichia coli* tRNA (Boehringer)] were added. Reactions were extracted twice with phenol/chloroform/3-methylbutan-1-ol (25:24:1, by vol.) and once with chloroform/3-methylbutan-1-ol (24:1, v/v). Reactions were then ethanol-precipitated for 1 h at -20 °C and spun in an Eppendorf centrifuge for 30 min at 4 °C. After the pellet had been washed with 70% (v/v) ethanol, the reaction mixtures were dried and dissolved in 4 μ l of RNA loading mix, denatured and separated on a 5% (w/v) polyacrylamide/urea gel. Gels were dried and exposed to Fuji X-ray films.

Footprinting analysis *in situ*

For footprinting reactions a *Hind*III/*Sac*I fragment, derived from P3-4TK-Luc encompassing P3-4 positions -195 to -163, was used as a probe. This fragment was end-labelled in the bottom strand with Klenow DNA polymerase and [α - 32 P]dCTP. EMSA reactions were performed with 10000 c.p.m. per reaction as described above. Subsequently, whole EMSA gels without prior identification of the retarded bands were used for footprinting experiments that were performed as described in [22]. After the DNA had been nicked by exposure for 10 min to the phenanthroline-copper ion solution, the bands were localized by autoradiography for 6 h and cut out of the gel. DNA fragments were then eluted overnight at 37 °C in 700 μ l of 0.5 M ammonium acetate/1 mM EDTA, precipitated with isopropanol, washed with 70% (v/v) ethanol, resuspended in 4 μ l of loading buffer, denatured and size-fractionated on a 10% (w/v) polyacrylamide/urea gel.

RESULTS

Transient transfections of HeLa and Hep3B cells with reporter constructs carrying the proximal IGF-II P3 region

To determine the role of the proximal P3 region in the regulation of human IGF-II P3 activity, transient transfection assays with P3 reporter constructs were performed in HeLa and Hep3B cells. Two different G-less cassette-luciferase constructs were tested for their promoter activity, the -289P3 construct (pFlag-Xba) containing the proximal promoter region between -289 and -7, and the -183P3 construct (pFlag-Sma) containing the region between -183 and -7. Transient transfection assays showed that the activity of the -289P3 construct in both cell

Table 1 Relative activities of the IGF-II P3 promoter constructs in HeLa and Hep3B cells

The activation of the P3 promoter constructs are reported relative to the activity of the -289P3 construct in HeLa cells, which was set at 100%. The activity of the empty vector pFlag was less than 3% of the activity of -289P3. Values are means $\pm$ S. D. for at least three independent experiments.

Promoter	Activity (%)	
	HeLa	Hep3B
-289P3	100 $\pm$ 13	86 $\pm$ 11
-183P3	32 $\pm$ 8	25 $\pm$ 6

Figure 1 DNase I footprinting analysis

DNase I footprinting analysis of the proximal promoter region (-290 to +134) with HeLa (left panel) and Hep3B (right panel) nuclear extracts. The protected region on the upper strand is indicated (-192 to -172). Symbols: +, nuclear extract added; -, no nuclear extract added.

lines was within the same range, indicating that the first 289 nt regulate the basal activity of this promoter and that the factors required for basal activity are present in both cell lines. Transfections with -183P3 resulted in a promoter activity that is approx. one-third that of -289P3 in both cell lines, indicating that the region between -289 and -183 contains elements important for basal P3 activity in HeLa and Hep3B cells (Table 1).

DNase I footprint analysis of the proximal P3 promoter region

The nuclear proteins that can bind to the promoter region between positions -289 and -150 were identified by DNase I footprinting analysis (Figure 1). In this experiment the region

Figure 2 Transcription-inhibition of proximal P3 constructs *in vitro* by oligonucleotides corresponding to the P3-4 element

(A) Transcription reactions with the -289P3 G-free cassette template (75 ng) were performed *in vitro* in the presence of a 0-, 75-, 150- or 225-fold molar excess of ds competitor oligonucleotides P3-4TOT, P3-1 and MCS respectively, with HeLa WCE. As an internal control, pAdML (200 ng) was included in the same reaction mixture. Lower panel: P3 transcription activity after quantification on a PhosphorImager. The P3 signal was normalized to the AdML signal. The activity of the empty vector pFlag was less than 1% of the activity of the -289P3 construct. (B) Transcription competition of the -289P3 construct (pFlag-Xba) and the -183P3 construct (pFlag-Sma) *in vitro* with a 0-, 75-, 150- or 225-fold molar excess of P3-4TOT oligonucleotides. Lower panel: P3 transcription activity after quantification on a PhosphorImager, after normalization.

located between -290 and +134 of P3 was used as probe and was incubated with HeLa or Hep3B nuclear extracts. Only one protected region was revealed between positions -289 and -150, designated P3-4, which is located in the upper strand between positions -192 and -172 relative to the transcription start site. In the more upstream region between -289 and -193 no further protected nucleotides were found (results not shown). The same footprint was found with HeLa as well as with Hep3B nuclear extracts (Figure 1), indicating that the P3-4 binding proteins are present in both cell lines.

Transcription-competition assays *in vitro* demonstrate the relevance of P3-4 for basal P3 activity

To explore the role of P3-4 for basal P3 activity, transcription assays were performed *in vitro* with HeLa WCEs. A ds oligonucleotide was synthesized (P3-4TOT) containing the protected region and some flanking nucleotides (positions -195 to -163). Transcription of the pFlag-Xba (-289P3) template was competed for by the addition of increasing amounts of the ds oligonucleotide P3-4TOT. To normalize the *in vitro* transcription assays, an AdML promoter construct was included in these

assays. For the competition studies two control ds oligonucleotides were included: a positive control containing the P3 CCAAT box (P3-1, positions -92 to -71), and a negative control consisting of an MCS. The importance of P3-1 for P3 activity has already been shown in previous experiments [12]. Figure 2(A) shows that transcription of the -289P3 template is efficiently inhibited by P3-4TOT and P3-1 oligonucleotides. The addition of an oligonucleotide containing the MCS does not affect transcription. Transcription of the AdML promoter is not affected by P3-4TOT, P3-1 or MCS oligonucleotides. This experiment showed that inhibition of transcription by P3-4TOT and P3-1 oligonucleotides of the -289P3 construct is specific, and that these regions are essential for -289P3 activity.

In agreement with the transient transfection assays described above, the promoter activity of the shorter pFlag-Sma construct (-183P3) in transcription assays *in vitro* was approx. two-fifths of the activity of the -289P3 construct (Figure 2B). This construct lacks the 5' region of the P3-4 element (-193 to -184) compared with the -289P3 construct. In contrast with the -289P3 template, transcription of the -183P3 template was not inhibited by the addition of increasing amounts of the P3-4TOT oligonucleotide, indicating that P3-4 is essential for activity of the -289P3 construct but not for -183P3 activity.

Figure 3 EMSA showing the presence of specific P3-4-binding proteins in HeLa and Hep3B cells

Left panel: EMSA with 2.5 µg of WCE from HeLa and Hep3B cells with the ds P3-4TOT probe. Right panel: EMSA-competition assay with 2.5 µg of HeLa WCE and P3-4TOT used as a ds DNA probe. The molar excess of unlabelled competitor oligonucleotides P3-4TOT and MCS is shown above the lanes. Protein-DNA complexes are numbered 1–5. Abbreviation: fp, free probe.

Element P3-4 is a binding site for multiple proteins

To identify the proteins binding to the P3-4 element, an EMSA was performed with an end-labelled ds P3-4TOT oligonucleotide and HeLa or Hep3B WCEs. Five different DNA-protein complexes, designated 1 to 5, were detected both in HeLa cells and in Hep3B cells (Figure 3, left panel). Whereas the expression levels of proteins 2, 3, 4 and 5 were comparable in the two cell lines, protein 1 was expressed at much higher levels in HeLa cells than in Hep3B cells. Because the P3-4 footprint is found with HeLa extracts as well as with Hep3B extracts and the transfection experiments showed a similar basal activity of the –289P3 construct in both cell lines, the difference in expression levels of protein 1 suggests that this protein is of minor importance for basal P3 activity.

To determine the specificity of all five complexes, EMSA competition experiments were performed with the unlabelled ds oligonucleotides P3-4TOT as a specific competitor and MCS as a non-specific competitor, with the use of HeLa extract (Figure 3, right panel). This experiment showed that the formation of complexes 1, 3, 4 and 5 was efficiently competed away by a 10-fold molar excess of the specific unlabelled P3-4TOT competitor oligonucleotide. Addition of the non-specific MCS did not affect the formation of any of these four complexes. In contrast, the formation of complex 2 was competed for with a very low efficiency by the specific as well as by the non-specific oligonucleotides, suggesting that complex 2 is non-specific. Taken together these results indicate that four different proteins (proteins 1, 3, 4 and 5) can bind specifically to the P3-4 element.

Two distinct protein-binding regions are present in the P3-4 element

To identify the binding regions of the four different proteins within the P3-4 element, ds oligonucleotides representing different regions of the entire element were synthesized (Figure 4A). Element P3-4 was divided into three partly overlapping regions, the left (P3-4L), middle (P3-4M) and right (P3-4R) regions of the complete element (P3-4TOT). Labelled P3-4TOT oligonucleotide was used as a probe; oligonucleotides containing these regions were used as competitors in an EMSA competition assay (Figure 4B). The four specific proteins 1, 3, 4 and 5 bound to the P3-4TOT oligonucleotide with high affinity. Complexes 1, 3, 4 and 5 were efficiently competed away with a 5-fold molar excess of unlabelled P3-4TOT competitor oligonucleotide. As observed before, protein 2 showed no specificity in oligonucleotide binding. The formation of a radiolabelled complex 2 was competed for with similar efficiency by the unlabelled P3-4TOT, P3-4L, P3-4M and P3-4R oligonucleotides, confirming that this complex is non-specific. Protein 1 showed a lower binding affinity for P3-4L, P3-4M and P3-4R than for P3-4TOT, indicating that none of these three oligonucleotides contains the complete binding region for protein 1. Protein 3 also bound all three subfragments with lower affinity. P3-4L was not bound at all, whereas P3-4M and P3-4R were bound with low affinity, indicating that protein 3 bound the middle and right parts of the P3-4 element. Proteins 4 and 5 bound P3-4M and P3-4R with reduced affinity, but bound P3-4L with high affinity, comparable to the binding of P3-4TOT. This indicates that proteins 4 and 5 have a binding site within the P3-4L region.

To determine the exact location of the binding sites of complexes 1, 3, 4 and 5, four different mutations were introduced into P3-4TOT and the corresponding oligonucleotides were used in EMSA competition assays with P3-4TOT as a probe (Figure 4C). For protein 1, P3-4TOTM1, P3-4TOTM2 and P3-4TOTM4 bound this protein as efficiently as the wild-type P3-4TOT oligonucleotide, whereas P3-4TOTM3 showed decreased binding of this protein. All four mutations had an effect on protein 3 binding, because this protein bound none of the mutant oligonucleotides with the same affinity as the wild-type P3-4TOT. Protein 3 binding was strongly decreased by the introduction of M1 and M3, because a 100-fold molar excess of P3-4TOTM1 and P3-4TOTM3 competitor oligonucleotide still gave a reasonable signal for protein 3. The most distinct effect of the introduction of a mutation was seen for proteins 4 and 5. Binding of these proteins was drastically decreased by the introduction of M1. The other mutations M2, M3 and M4 had no effect on the binding affinity of these two proteins, suggesting that the binding site of proteins 4 and 5 is located in the region where M1 is introduced. In good agreement with the previous results this binding site is located within the left region of the P3-4 element.

From the results described it was still not clear to which region within the P3-4 element protein 3 could bind. For further analysis, HeLa WCE was partly fractionated over a phosphocellulose column with a continuous salt gradient from 0.04 to 1 M KCl. Different fractions were obtained containing protein 3. One peak fraction containing protein 3 was selected and tested in a competition EMSA for its binding affinity for the different P3-4 mutants (Figure 4D). This experiment showed that M1 and M2 have a marginal effect on protein 3 binding. M3 severely decreased protein 3 binding, whereas protein 3 also bound P3-4TOTM4 with a lower affinity. Because mutations M3 and M4 are located close to each other, this suggests that protein 3 binds the region of these mutations, with a high-affinity binding site in the region where M3 is located.

Figure 4 EMSA competition with oligonucleotides containing different truncations and mutations in the P3-4 element

(A) Nucleotide sequences of the different oligonucleotides used in EMSA competition analysis. Mutations are indicated by bold letters. (B) EMSA competition with truncated P3-4 oligonucleotides, with 2.5 μ g of HeLa WCE and P3-4TOT as a probe. In the reaction, binding was competed for with a 0-, 2.5-, 5-, 10-, 25-, 100- or 250-fold molar excess of unlabelled ds P3-4TOT, P3-4L, P3-4M or P3-4R oligonucleotides. The competitor used is indicated above each panel; the protein-DNA complexes 1–5 are indicated at the left. (C) EMSA competition with mutated P3-4 oligonucleotides, with 2.5 μ g of HeLa WCE and P3-4TOT as a probe. Binding was competed for by the addition of a 0-, 2.5-, 5-, 10-, 25-, 100- or 250-fold molar excess of unlabelled ds oligonucleotides P3-4TOT, P3-4TOTM1, P3-4TOTM2, P3-4TOTM3 and P3-4TOTM4. (D) EMSA competition as in (C), but 0.5 μ g of a phosphocellulose fraction containing protein 3 was used as protein source.

Figure 5 Effects of oligonucleotides containing P3-4 mutations and truncations on transcription of the $-289P3$ construct

Upper panel: transcription-competition of $-289P3$ activity *in vitro* by the addition of a 0-, 50-, 100-, 150- or 200-fold molar excess of oligonucleotides containing different P3-4 mutations. The activities of $-289P3$ after competition with oligonucleotides P3-4TOT, P3-4TOTM1, P3-4TOTM2, P3-4TOTM3 and P3-4TOTM4 are shown. The promoter activity of P3 without the addition of competitor oligonucleotide was set at 100%. All signals were normalized to the AdML promoter signal. Lower panel: transcription-competition of $-289P3$ activity *in vitro* by the addition of a 0-, 50-, 100-, 150- or 200-fold molar excess of competitor oligonucleotides. These competitor oligonucleotides contained different truncations of the P3-4TOT oligonucleotide as illustrated in Figure 4. The P3 signal was normalized to the internal control AdML.

In conclusion, the P3-4 element can be subdivided into two distinct regions. Proteins 4 and 5 bind the left region containing a poly(dG)·poly(dC) region where M1 is located; protein 3 binds the middle region where M3 is located.

Relevance of the two P3-4 protein-binding regions for P3 regulation

To determine the relative importance of the two different protein-binding regions for P3 activity, transcription-competition assays were performed *in vitro* with different P3-4 mutant ds oligonucleotides as competitors. Increasing amounts of competitor oligonucleotide were incubated with HeLa transcription extract, after which transcription reactions were performed *in vitro* with the $-289P3$ template (Figure 5, upper panel). All four P3-4 mutant oligonucleotides inhibited transcription of the $-289P3$ template, whereas the internal control, the AdML template, was not affected. P3-4TOTM2, P3-4TOTM3 and P3-4TOTM4 oligonucleotides inhibited transcription with the same efficiency as the wild-type P3-4TOT did, resulting in a maximal inhibition of 60% with a 200-fold molar excess of competitor oligonucleotide.

Figure 6 Copper-phenanthroline footprinting of P3-4-binding proteins

Upper panel: the different P3-4-binding complexes were separated by EMSA, subjected to the footprinting reaction as described in the Materials and methods section, and located by autoradiography. Subsequently, the regions containing complexes 3, 4 and 5 and the free probe were cut out of the gel, then DNA was eluted and size-separated. Abbreviations: GA, marker lane; fp, free probe. The DNA sequences of the protected boxes A and B are depicted at the left. Lower panel: schematic overview of the protected regions detected by copper-phenanthroline footprinting on the lower strand (box A and box B) and by DNase I footprinting on the upper strand.

Addition of the P3-4TOTM1 oligonucleotide resulted in a maximal inhibition of only 25% with a 200-fold molar excess of P3-4TOTM1 competitor. The decrease in capacity of P3-4TOTM1 to inhibit transcription of the -289P3 template suggests that the mutated region is important for P3 activity. From EMSA experiments described above it is obvious that the M1 mutation drastically decreases the binding of proteins 4 and 5 to the P3-4 element, suggesting that these proteins are important for P3 activity.

Transcription reactions were also performed *in vitro* in the presence of increasing amounts of oligonucleotides containing different P3-4 regions: P3-4L, P3-4M and P3-4R (Figure 5, lower panel). The addition of increasing amounts of P3-4M and P3-4R competitor resulted in a very moderate inhibition (at most 15%) with a 200-fold molar excess of competitor oligonucleotide. In contrast, the addition of increasing amounts of P3-4L oligonucleotide resulted in a 55% maximal inhibition of transcription with a 200-fold molar excess of this competitor, comparable to the inhibition with the P3-4TOT oligonucleotide itself. The

Figure 7 Box A-binding proteins need Zn^{2+} ions for their DNA-binding activity

EMSA showing the requirement of Zn^{2+} ions for the DNA-binding activity of proteins 4 and 5. HeLa WCE was preincubated for 15 min with 1 mM EDTA (lane 2), after which 0.5 (lane 3) or 1 mM (lane 4) ZnCl_2 was added during the binding reaction. Protein-DNA complexes were numbered 1-5 as indicated at the left.

EMSA experiments described above reveal that proteins 4 and 5 have a binding site within P3-4L, whereas protein 3 does not bind P3-4L. This suggests that the high-efficiency inhibition of transcription of this oligonucleotide is the consequence of competition for proteins 4 and 5, but not for protein 3. In conclusion, two lines of evidence indicate that proteins 4 and 5 are essential proteins for P3 activity.

Differential footprint analysis of the P3-4-binding proteins

For further characterization of the protein-binding sites within P3-4, copper-phenanthroline footprinting was performed. This technique allows the identification of the binding regions of the individual proteins. Complex 3 and the doublet of complexes 4 and 5 were separated by EMSA, subjected to the footprinting reaction and separately isolated from the gel. Subsequent gel electrophoresis of the eluted bands revealed that two different regions on the lower strand, designated A and B, are protected by these proteins. Region A lies on the lower strand between -193 and -188 relative to the transcription start site and was protected by proteins 4 and 5. Protein 3 protected a region between -183 and -172 on the lower strand, designated B (Figure 6, upper panel). This indicates that P3-4 can be subdivided into two non-overlapping boxes, A (-193 to -188) and B (-183 to -172), bound by proteins 4/5 and 3 respectively. Box A represents the sequence 5'-188GGGGGG-193-3', present in a region of 11 dG residues (-194 to -184); box B represents the sequence 5'-172CCCTGGGGCCCC-183-3'. These two boxes are separated by four nucleotides. Interestingly, the combined region protected by proteins 3, 4 and 5 (-193 to -172) on the lower strand completely covers the initial DNase I footprint (Figure 1) that was found on the upper strand (-192 to -172), obtained with HeLa and Hep3B nuclear extracts (Figure 6, lower panel).

This suggests that proteins 3, 4 and 5 can bind the P3-4 element simultaneously without any steric hindrance, thereby also protecting the four nucleotides (–187 to –184) that separate boxes A and B.

Box A-binding proteins need Zn^{2+} ions for their binding activity

To characterize further the proteins binding P3-4, the metal ion requirement for their DNA-binding activity was investigated. HeLa WCE was preincubated for 15 min with the bivalent-chelating agent EDTA (Figure 7). Preincubation of the extract with EDTA at a concentration of 0.5 mM inhibited the formation of complexes 4 and 5, whereas the formation of the other complexes was not affected. Reconstitution experiments were performed to establish the specificity of the metal ion requirement. The addition of Zn^{2+} ions completely restored the binding activity of proteins 4 and 5 at concentrations of 0.5 and 1 mM, whereas the addition of equimolar amounts of Mg^{2+} could not restore binding (results not shown). This suggests that the box A-binding proteins 4 and 5 need zinc for their binding, suggesting that these factors might be zinc-finger transcription factors.

DISCUSSION

The IGF-II promoter P3 has been shown to contain a functional TATA and CCAAT box. Several other regulatory elements, such as binding sites for the transcription factor EGR-1, have been identified in the proximal region. Transient transfection experiments with truncated promoter P3 constructs revealed that the distal promoter region (–1300 to –289) is responsible for the cell-type-specific expression of IGF-II. The proximal region (–289 to +140) exhibits basal levels of expression that are similar in several different cell lines, including HeLa and Hep3B.

In the present paper we show that the region between –289 and –183 is essential for basal P3 activity. DNase I footprinting analysis of the region between –289 and –183 revealed the presence of a single protected element located at positions –192 to –172, designated P3-4. EMSA experiments of P3-4 with extracts of both HeLa and Hep3B cells further indicated that several proteins bind to P3-4. The importance of P3-4 was confirmed by transcription experiments *in vitro* where competition with oligonucleotides containing this P3-4 sequence resulted in the efficient inhibition of transcription of –289P3 but not –183P3, establishing a key role for P3-4 in –289P3 activity.

Initial characterization of the proteins binding to this element has resulted in the identification of four specific proteins that are present in HeLa and Hep3B cells. The expression levels of proteins 3, 4 and 5 are comparable in both cell lines. Because the expression level of protein 1 is much higher in HeLa cells than in Hep3B cells and because activity level of –289P3 is comparable in both cell lines, we conclude that this protein is not involved in basal P3 activity. Because EMSA competition analysis also showed that complex 2 is non-specific, we further focused on proteins 3, 4 and 5 for their involvement in P3 regulation.

The usage of oligonucleotides containing different mutations in the P3-4 region in EMSA competition assays revealed the presence of two important regions for protein binding. Footprinting experiments of the different proteins have identified two non-overlapping boxes, designated box A (–193 to –188) and box B (–183 to –172), on the lower strand. Protein 3 binds box B; proteins 4 and 5 bind box A. Because the complexes containing proteins 4 and 5 ran very close together in EMSA and their behaviour was the same in EMSA competition experiments, the

two proteins cannot be discriminated. In EMSA competition experiments with WCE, it was difficult to show the effect of the introduction of a mutation in box B (M3) on the binding of protein 3. However, when using a more purified protein fraction, partly separated on a phosphocellulose column, decreased binding of protein 3 to an oligonucleotide containing M3 could be demonstrated. Efficient competition for protein 3 might be inhibited by the high protein concentration in the WCE, an effect that is circumvented by the use of a more extensively purified fraction. The availability of oligonucleotides containing mutations in the two boxes allowed us to establish the importance of these boxes for P3 activity. P3-4TOTM3 (mutated in box B) was able to inhibit transcription from the –289P3 template with the same efficiency as did P3-4TOT (60% maximal inhibition). This suggests that protein 3 is not essential for P3 activity in these transcription experiments *in vitro*.

The addition of P3-4TOTM1 oligonucleotides (mutated in box A) had a very moderate effect on transcription (25% maximal inhibition), suggesting that proteins 4 and 5, which bind the M1 region, are indeed important for –289P3 activity. This was confirmed by the inhibition of transcription by the P3-4L oligonucleotide (55% maximal inhibition). Both proteins 4 and 5 bound this oligonucleotide with the same high affinity as P3-4TOT. This is also in good agreement with the transfection results, where deletion of the left P3-4 region (containing box A) in the –183P3 construct resulted in a decrease in activity to one-third compared with the –289P3 construct. The right P3-4 region, containing box B, is still present in both constructs. Taken together, these results lead us to conclude that box A, bound by proteins 4 and 5, is essential for P3 activity.

Binding of proteins to poly(dG)·poly(dC) stretches has been reported for the so-called G-string-binding factors. In the chicken, two G-string-element-containing promoters have been described that are bound by the chicken erythrocyte protein called β -globin protein 1, BGP 1 [23,24]. In the sea urchin (*Psammechinus miliaris*) the purification of another G-string-binding protein of 59.5 kDa, called sea-urchin G-string factor 1 (suGF1), has been described [25,26]. The promoter of the cell-lineage-specific gene *LpS- β* of the sea urchin *Lytechinus pictus* also contains a G-string element dG₆dC that can be bound by suGF1 [27]. The fact that proteins 4 and 5 share some characteristics with the above-mentioned G-string proteins suggests that proteins 4 and 5 are members of the family of G-string factors. This assumption is based on the following characteristics. First, the proteins bind to an 11 bp (dG)-homopolymer sequence in the IGF-II P3 promoter region, of which six contiguous dG residues (box A) are directly interacting with proteins 4 and 5. Secondly, the Zn^{2+} requirement for DNA binding of proteins 4 and 5 suggests that these two proteins could be zinc-finger transcription factors. Interestingly, the requirement for bivalent ions has also been shown for the G-string-binding factor BGP1, which needs Zn^{2+} ions for its DNA-binding activity [23]. Furthermore it was shown that the sea-urchin G-string-binding factor suGF1 is also tightly complexed with a cation [26]. Thirdly, proteins 4 and 5 are unable to bind to single-stranded DNA (results not shown). This ds-DNA-binding specificity was also shown to exist for the G-string-binding factor suGF1 [26].

The distinct non-overlapping binding sites for the different proteins 3, 4 and 5, spaced by four nucleotides, suggests that the proteins bind the DNA simultaneously without steric hindrance. The initial DNase I footprint (Table 1) confirmed this finding, by a footprint that spanned the complete region containing boxes A and B in HeLa and Hep3B cells. Here we describe the separation of the P3-4 element into two distinct protein-binding regions. From these results, no evidence for physical interactions between

the box A-binding and box B-binding proteins was found. Further purification will be necessary for such studies.

Several lines of evidence indicate that the ubiquitous G/C box-binding factor Sp1 is not involved in P3 regulation via P3-4. First, the P3-4 element does not contain a consensus Sp1-binding site and in EMSA experiments it was shown that Sp1 is not able to bind the P3-4 element (results not shown). Secondly, none of the four specific HeLa complexes 1, 3, 4 or 5 competes with an oligonucleotide containing the Sp1 consensus binding site (results not shown). Thirdly, the position of the specific complexes differs for all proteins from the position of Sp1 in EMSA experiments.

In conclusion, functional analysis of the P3-4 element has revealed two binding regions, each able to bind specific proteins. Box A, containing a G-string, binds two proteins essential for the basal activity of P3. Future studies will focus on the isolation and characterization of the P3-4-binding proteins and the determination of their transcriptional activity.

We thank R. Schiphof for providing us with HeLa cells. This work was supported by a grant from the Netherlands Organization for the Advancement of Pure Research (N.W.O.).

REFERENCES

- Daughaday, W. H. and Rotwein, P. (1989) *Endocr. Rev.* **10**, 68–86
- DeChiara, T. M., Efstradiatis, A. and Robertson, E. J. (1990) *Nature (London)* **345**, 78–80
- Recio-Pinto, E., Lang, F. L. and Ishii, D. N. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2562–2566
- Ewton, D. Z. and Florini, J. R. (1981) *Dev. Biol.* **56**, 31–39
- Cohick, W. S. and Clemmons, D. R. (1993) *Annu. Rev. Physiol.* **55**, 131–153
- de Pagter-Holthuisen, P., Jansen, M., van Schaik, F. M. A., van der Kammen, R., Oosterwijk, C., Van der Brande, J. L. and Sussenbach, J. S. (1987) *FEBS Lett.* **214**, 259–264
- de Pagter-Holthuisen, P., Jansen, M., van der Kammen, R., van Schaik, F. M. A. and Sussenbach, J. S. (1988) *Biochim. Biophys. Acta* **950**, 282–295
- Holthuisen, P., van der Lee, F. M., Ikejiri, K., Yamamoto, M. and Sussenbach, J. S. (1990) *Biochim. Biophys. Acta* **1087**, 341–343
- van Dijk, M. A., van Schaik, F. M. A., Bootsma, H. J., Holthuisen, P. and Sussenbach, J. S. (1991) *Mol. Cell. Endocrinol.* **81**, 81–94
- Drummond, I. A., Madden, S. L., Rohwer-Nutter, P., Bell, G. I., Sukhatme, V. P. and Rauscher, III, F. J. (1992) *Science* **257**, 674–677
- Madden, S. L., Cook, D. M., Morris, J. F., Gashler, A., Sukhatme, V. P. and Rauscher, III, F. J. (1991) *Science* **253**, 1550–1553
- van Dijk, M. A., Holthuisen, P. and Sussenbach, J. S. (1992) *Mol. Cell. Endocrinol.* **88**, 175–185
- Sawadogo, M. and Roeder, R. G. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4394–4398
- Knowles, B. B., Howe, C. C. and Aden, D. P. (1980) *Science* **209**, 497–499
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- deWet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. and Subramani, S. (1987) *Mol. Cell. Biol.* **7**, 725–737
- Pruijn, G. J., van Driel, W., van Miltenburg, R. T. and van der Vliet, P. C. (1987) *EMBO J.* **6**, 3771–3778
- Manley, J. L., Fire, A., Samuels, M. and Sharp, P. A. (1983) *Methods Enzymol.* **101**, 568–582
- Samuels, M., Fire, A. and Sharp, P. A. (1982) *J. Biol. Chem.* **257**, 14419–14427
- Pruijn, G. J., van Driel, W. and van der Vliet, P. C. (1986) *Nature (London)* **322**, 656–659
- Shen, R. F., Clift, S. M., DeMayo, R. N., Sifers, R. N., Finegold, M. J. and Woo, S. L. C. (1989) *DNA* **8**, 101–108
- Kuwabara, M. D. and Sigman, D. S. (1987) *Biochemistry* **26**, 7234–7238
- Lewis, C. D., Clark, S. P., Felsenfeld, G. and Gould, H. (1988) *Genes Dev.* **2**, 863–873
- Clark, S. P., Lewis, C. D. and Felsenfeld, G. (1990) *Nucleic Acids Res.* **18**, 5119–5126
- Hapgood, J. and Patterson, D. (1994) *Mol. Cell. Biol.* **14**, 1402–1409
- Patterson, D. and Hapgood, J. (1994) *Mol. Cell. Biol.* **14**, 1410–1418
- Seid, C. A., Sater, A. K., Falzone, R. L. and Tomlinson, C. R. (1996) *DNA Cell Biol.* **15**, 511–517