

# The Pal elements in the upstream glucokinase promoter exhibit dyad symmetry and display cell-specific enhancer activity when multimerised

J. M. Moates<sup>1, 2</sup> · M. A. Magnuson<sup>3</sup>

<sup>1</sup> Department of Medicine, Vanderbilt University School of Medicine and The Nashville Veterans' Affairs Medical Center, Nashville, Tennessee, USA

<sup>2</sup> Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Alabama at Birmingham and The Birmingham Veterans' Affairs Medical Center, Birmingham, AL, USA

<sup>3</sup> Departments of Molecular Physiology and Biophysics and Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

## Abstract

**Aims/hypothesis.** The upstream glucokinase ( $\beta$ GK) promoter has previously been shown to contain two 9-bp sequences, termed the Pal motifs, that are conserved in humans, rats and mice. These motifs are necessary for expression of the  $\beta$ GK promoter in pancreatic beta cell lines and pituitary corticotrope cell lines. DNA probes containing the Pal motifs bind cell-type-specific protein complexes, but these motifs have not been completely characterised.

**Methods.** Methylation interference and ultraviolet-crosslinking analysis were utilised to characterise, at the single nucleotide level, sites of protein binding within the elements themselves. To determine the function of these elements, mutational analysis of the  $\beta$ GK promoter and of multimerised GK promoter sequences was performed.

**Results.** Both Pal elements are 14 bp in length and have dyad symmetry. However, while the Pal-1 element is a perfect inverted repeat (GTCACCA-TGGT-GAC), the Pal-2 element (GTCACCA-TAGAAAC) is

an imperfect repeat. Ultraviolet-crosslinking analysis using nuclear extracts prepared from hamster insulinoma tumour (HIT) cells revealed that the three resolvable complexes that bind to the Pal-1 and Pal-2 elements contain different ratios of three proteins of different size (~90, 110 and 150 M<sub>r</sub>). Mutation of a single nucleotide binding site abrogates  $\beta$ GK promoter activity. Multimerised repeats of the Pal-1 element augment transcription in HIT cells, but not in baby hamster kidney (BHK) cells.

**Conclusions/interpretation.** These results suggest that different combinations of three proteins of different size bind to the Pal elements, probably as homodimers and heterodimers. Together, these results indicate that the  $\beta$ GK promoter contains two novel 14-bp elements that, when multimerised, exhibit enhancer activity specific to neuroendocrine cells.

**Keywords** Dimer · Dyad symmetry · Enhancer · Glucokinase · Inverted repeat · Methylation interference · Promoter

Received: 22 March 2004 / Accepted: 2 June 2004

Published online: 10 September 2004

© Springer-Verlag 2004

J. M. Moates (✉)

Division of Endocrinology, Diabetes and Metabolism,  
Department of Medicine,  
University of Alabama at Birmingham  
and The Birmingham Veterans' Affairs Medical Center,  
Rm 706 BDB, 1808 7th Ave. So., Birmingham,  
AL 35294-0012, USA

E-mail: mmoates@uab.edu

Tel.: +1-205-934-9819, Fax: +1-205-934-4389

## Introduction

The upstream glucokinase ( $\beta$ GK) promoter is expressed in neuroendocrine cells of the pancreas, pituitary, gut and brain [1, 2, 3]. Previous fusion gene studies in permanent cell lines and transgenic mice have established that a 294-bp region of the  $\beta$ GK promoter con-

**Abbreviations:** BHK, baby hamster kidney · BrdU, 5-bromo-2'-deoxyuridine · EMSA, electrophoretic mobility-shift assay · GK, glucokinase · HIT, hamster insulinoma tumour · PEBP, pal-element-binding protein · UV, ultraviolet

fers neuroendocrine-cell-specific expression of  $\beta$ GK in pancreatic beta cells and pituitary corticotropes [1, 4, 5]. Mutagenesis and DNA–protein-binding analysis of this region of the promoter have shown that several distinct sequence motifs are important for transcription of the  $\beta$ GK promoter in neuroendocrine cells. The first of these motifs are the upstream promoter elements, which are conserved AT-rich sequences that are bound by the homeodomain transcription factor Pdx-1 [6]. However, Pdx-1 is primarily restricted to islet beta cells [7] and  $\beta$ GK gene expression extends to a broader set of neuroendocrine cells, indicating that additional factors are required for regulating  $\beta$ GK gene transcription in these cells.

A second motif in the GK promoter is a prototypical E-box motif at position –221 to –216 which is bound by the transcription factor BETA2/NeuroD1 [8]. BETA2 is a basic helix-loop-helix transcription factor that appears to regulate the transcription of a diverse group of genes, including the preproglucagon gene in alpha cells [9], the *POMC* gene in pituitary corticotropes [10] and the secretin gene in gut enteroendocrine cells [11]. However, BETA2 is more broadly expressed in the brain in cells lacking  $\beta$ GK [12], suggesting that additional factors are involved that may restrict  $\beta$ GK expression to a discrete subset of cells.

Another pair of sequence motifs in the  $\beta$ GK promoter that are essential for  $\beta$ GK transcription in neuroendocrine cells consists of two 9-bp repeats termed the Pal motifs [4, 5]. The precise sequences of the elements that contain the Pal motifs have not been determined, and the proteins that bind to them are yet to be identified. Promoter mutations that disrupt the structure of the Pal motif markedly reduce  $\beta$ GK promoter expression in neuroendocrine cells [5]. When DNA probes containing the Pal motif sequences are used in electrophoretic mobility-shift assays (EMSAs), DNA–protein complexes are observed that are cell-type-specific. The mobility and abundance of these DNA–protein complexes are dependent upon whether the cell extracts were prepared from cells that express, or do not express  $\beta$ GK. Specifically, the banding patterns observed using nuclear extracts from two neuroendocrine cell types that express  $\beta$ GK (hamster insulinoma tumour [HIT] M2.2.2 cells and AtT-20 cells) are different to those obtained using extracts from two cell types that do not express  $\beta$ GK (NIH-3T3 cells and CV-1 cells) [5]. The complexity of the DNA–protein complexes that formed on probes containing the Pal motifs, as well as the lack of clear knowledge about the actual DNA element involved, prompted us to perform an analysis of the nucleotides that contribute to the formation of the different DNA–protein complexes observed.

This study aimed to clearly establish protein–DNA contact points in  $\beta$ GK-expressing and non-expressing cell types using methylation interference analysis.

## Materials and methods

**Plasmid DNAs.** Single bp mutations in Pal-1 or Pal-2 were made using a Quick-change mutagenesis kit (Stratagene, LaJolla, Calif., USA) in a plasmid (p $\beta$ GK-luc) containing wild-type rat  $\beta$ GK promoter sequences (–402 to +14) linked to a luciferase reporter gene. All mutations were verified by DNA sequencing. The minimal Pal-element–reporter constructs were made by ligating complementary oligonucleotides containing three tandem copies of the Pal-element-binding motif. The top strand DNA sequence was: 5'-TCGAGGTCACCATGGT-GACAGGTCACCATGGT-GACAGGTCACCATGGT-GACAG. The underlined "C" was changed to an "A" in the mutated form of the oligonucleotide used in the enhancer constructs. Oligonucleotides were gel-purified, treated with polynucleotide kinase and ATP, and annealed prior to ligation between the *Xho*I and *Sal*I sites upstream of a minimal SV40 promoter in the plasmid pGL2 (Promega, Madison, Wis., USA). The mutant Pal elements contained a C to A substitution (G to T in the complementary strand) at position 3 in the left half of the element. The orientations of the wild-type and mutant tandem Pal elements within the plasmid were determined by DNA sequencing. Plasmid DNAs were purified through two CsCl density gradients prior to transfection.

**Cell culture and transfection.** HIT M2.2.2 and baby hamster kidney (BHK) cells were grown to 75% confluence in Dulbecco's modified essential medium supplemented with 10% FCS (BHK cells) or 2% FCS / 15% horse serum (HIT cells) with streptomycin and penicillin (50  $\mu$ g/ml each). Cells were plated on to tissue culture dishes at a density of  $1 \times 10^6$  cells per 75 cm<sup>2</sup> and transfected using the CaPO<sub>4</sub> co-precipitation method as described previously [4]. Transfected cells were harvested 40 h later and assayed for luciferase activity as described previously [13]. Transfection efficiency was normalised using a  $\beta$ -galactosidase expression vector as described previously [4].

**Electrophoretic mobility-shift assays.** Gel purification, annealing and purification of radiolabelled probes for the EMSAs were performed as described previously [4]. Oligonucleotide pairs were end-labelled using [ $\gamma$ -<sup>32</sup>P]ATP and T4 polynucleotide kinase. Nuclear extracts were prepared by the method described by Shapiro et al. [14] with the modifications described by Shelton et al. [4] and stored in aliquots at –70 °C until used. Gel-shift reactions (20  $\mu$ l) contained 5  $\mu$ g of nuclear extract, 20 mmol/l MgCl<sub>2</sub>, 1  $\mu$ g of polydeoxyinosinic-deoxycytidylic acid, 1  $\mu$ g of sheared denatured salmon sperm DNA, 12.5 mmol/l HEPES (pH 7.8), 2.5 mmol/l dithiothreitol, 0.1 mmol/l EDTA, 5% glycerol and 50,000 cpm of radiolabelled oligonucleotide. After incubation at 4 °C for 20 min, the reaction products were separated by electrophoresis through a 4% acrylamide gel buffered with 250 mmol/l Tris, 1.9 mmol/l glycine and 10 mmol/l EDTA at 4 °C. The gels were then dried and an autoradiograph was obtained. Oligonucleotide competitors contained A to C or G to T point mutations.

**Methylation interference assay.** The top and bottom strands of the oligonucleotide duplexes were gel-purified, end-labelled as above and annealed to a 2-fold excess of an unlabelled complementary oligonucleotide. Probe preparation and cleavage was essentially as described previously [15], with the following modifications. Oligonucleotide probes were hemimethylated in a 200- $\mu$ l reaction containing 50 mmol/l sodium cacodylate, 1 mmol/l EDTA and 1  $\mu$ l of dimethyl sulphate (Sigma-Aldrich, St. Louis, Mo., USA) for 120 s at 24 °C. Reactions were terminated by the addition of 50  $\mu$ l of stop buffer containing 1.5 mol/l NaOAc and 1 mol/l  $\beta$ -mercaptoethanol. Methyl-

ated probes were separated from unincorporated radionucleotide on a G-50 Sephadex column equilibrated with TE (10 mmol/l Tris Base, 1 mmol/l EDTA, pH 8.0). A total of  $5 \times 10^5$  cpm of probe were incubated with 10  $\mu$ g of HIT or BHK cell nuclear extract at 4 °C for 20 min in a typical electrophoretic mobility-shift reaction. DNA-protein complexes were separated from unbound DNA by electrophoresis through a 4% polyacrylamide gel at 4 °C. The gel was then wrapped in cellophane and an autoradiograph was obtained. Gel slices containing bound and free probe were cast into 1.2% agarose gels, electro-blotted on to DEAE membranes and then eluted at 65 °C with 350  $\mu$ l of high-salt buffer containing 50 mmol/l Tris-HCl, 1 mol/l NaCl and 10 mmol/l EDTA. After two extractions with phenol/chloroform, the probe was precipitated by the addition of ice-cold ethanol and 25  $\mu$ g of oyster glycogen. After centrifugation at 14,000 rev/min for 15 min, the precipitated pellet was resuspended in 43  $\mu$ l of 10 mmol/l NaPO<sub>4</sub> and 7.5  $\mu$ l of 1 mol/l NaOH, then incubated for 30 min at 90 °C. After the addition of 7.0  $\mu$ l of 1 mol/l HCl, the cleaved fragments were precipitated as above, resuspended in 98% formamide/10 mmol/l EDTA, and equal counts of bound and free products were separated by electrophoresis through a 20% denaturing polyacrylamide gel, which was subsequently subjected to autoradiography.

**Ultraviolet-crosslinking studies.** A Pal-1 probe was synthesised that contained 5-bromo-2'-deoxyuridine (BrdU) close to critical binding sites in both the top (5'-ACAGACTGGUCAC-CAUGGUGACAG-3') and bottom (5'-CTGTCACCAUGGU-GACCACTCTGT-3') DNA strands [16]. The reaction was a typical gel-shift reaction except that 200,000 cpm of the radiolabelled Pal-1-BrdU probe was combined with 40  $\mu$ g of HIT or BHK nuclear extract in a 40- $\mu$ l reaction. After a 20-min incubation period at 4 °C, the reaction was placed in a single well of a 96-well tissue culture dish, wrapped in cellophane and placed on a bed of ice 2 cm below the ultraviolet (UV) lamps in a Stratilinker (Stratagene) and irradiated for 20 min at maximum energy. The products were then separated by electrophoresis through either a 7.5% SDS polyacrylamide gel for 5 h at 30 mA or a 6% electrophoretic mobility-shift gel for 5 h at 165 V. The electrophoretic mobility-shift gel was wrapped in cellophane and an autoradiograph was obtained. After this, the small gel slices containing DNA-protein complexes were excised, stuffed into the well of a 4.5% stacking gel and separated by electrophoresis through a 7.5% SDS polyacrylamide gel for 5 h at 30 mA. Autoradiographs of dried gels were subsequently obtained.

**Statistical analysis.** The mean values for the activity of the promoter fragments were compared using the Student's *t* test, and determined to be statistically significant at a *p* value of less than 0.05. The *p* values are listed in the figure legends.

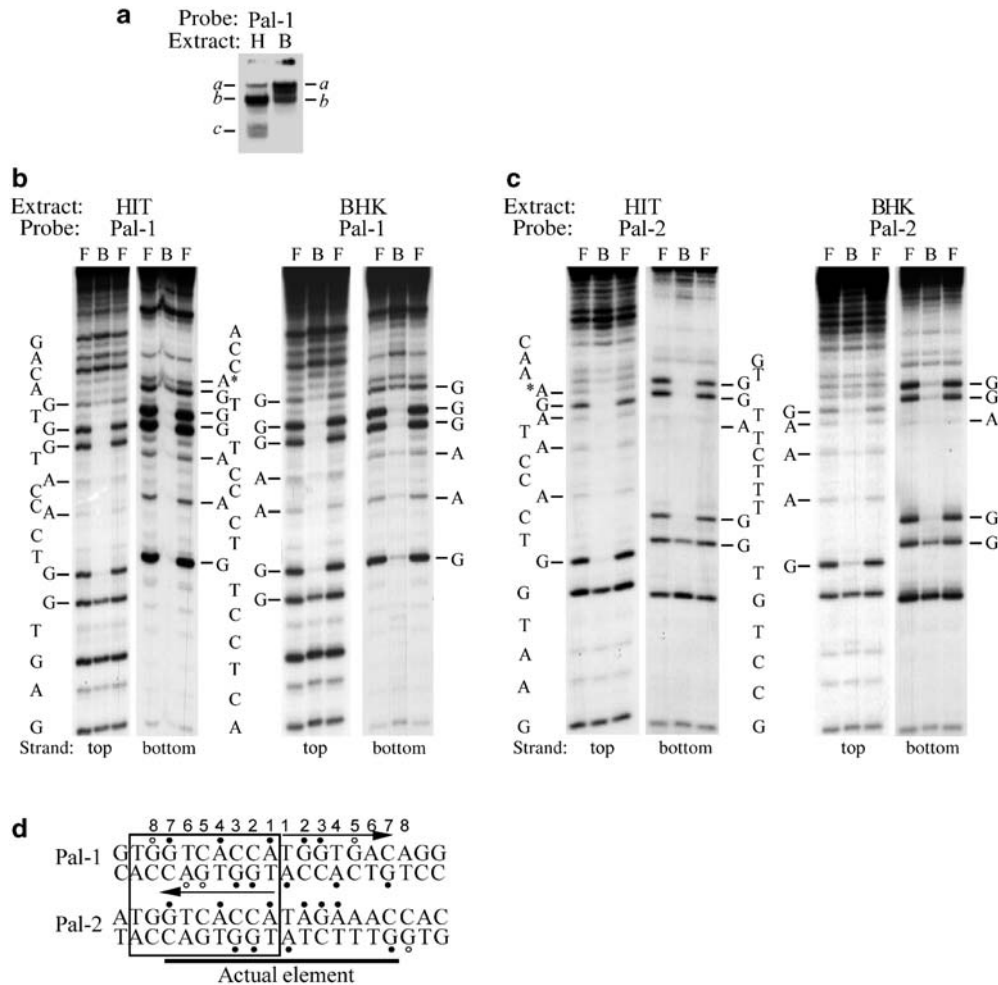
## Results

**Contact points for the Pal-element-binding proteins.** Previous EMSAs of probes containing Pal motif sequences and nuclear proteins from both  $\beta$ GK-expressing and non-expressing cells have shown that multiple DNA-protein complexes are formed [5]. We used BHK fibroblasts as a non-expressing cell type since they do not express GK mRNA (as assessed by RT-PCR, data not shown) and, like HIT cells, are derived from Syrian hamsters. The complexes formed using BHK cell extracts and a Pal-1 probe are similar to

those previously observed using NIH-3T3 cells, but differ from those produced using HIT cell extracts with respect to mobility and relative abundance (Fig. 1a). The Pal-2 probe gave the same binding pattern as Pal-1 on incubation with HIT and BHK cell extracts (data not shown). To determine the precise contact points between the Pal-element-binding proteins (PEBPs) and the DNA probes, a DNA methylation interference analysis was performed for the three complexes from the HIT cell extract (denoted *a*, *b* and *c*) and the two complexes from the BHK cell extract (*a* and *b*). The probes used contained either a Pal-1 or Pal-2 motif and formed the desired DNA complexes in EMSAs. Since similar results were obtained in all cases, only the results for the most abundant complexes (*b* for HIT cells and *a* for BHK cells) and those for the top and bottom strands of the Pal-1 and Pal-2 probes are shown in Figures 1b and 1c respectively.

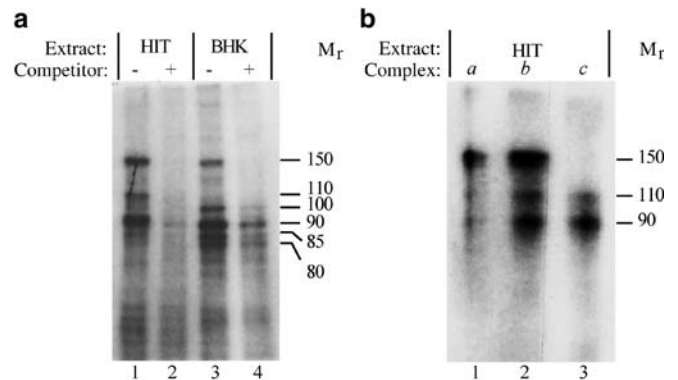
Similar contact points were observed for both the HIT and BHK nuclear extracts, except for an adenine nucleotide that was contacted in HIT cells, but not in BHK cells (indicated by the asterisk in Figure 1b). Further inspection of the autoradiograms revealed that proteins bound to an element that had perfect dyad symmetry (GTCACCA-TGGTGAC) when incubated with the Pal-2 probe. In contrast, the contact points on the Pal-2 probe revealed a similar, but imperfect palindromic repeat (GTCACCA-TAGAAAC). The sequences and contact points are illustrated in Figure 1d. The left half of the palindromic DNA sequence is the same for both Pal-1 and Pal-2, and encompasses seven nucleotides of the originally described Pal motif. In contrast, the right half of the 14-bp region of DNA-protein interaction differed between Pal-1 and Pal-2 at several nucleotides (Fig. 1d).

**Ultraviolet-crosslinking studies of the Pal-element-binding proteins.** UV-crosslinking experiments were performed to identify any differences between the complexes (shown in Fig. 1a) that bind to the Pal elements in HIT and BHK cells. We initially tested whether a Pal-1 probe containing BrdU at several thymidines within the element formed the same complexes as were previously detected by UV-laser crosslinking [5]. As before, bands of 150, 110, and 90 M<sub>r</sub> were observed using a HIT cell extract (Fig. 2a). The sizes of the bands observed in BHK cells (Fig. 2a) were similar, but not identical. For both cell types, 150-M<sub>r</sub> bands were observed, as well as a 100-M<sub>r</sub> protein that was much more abundant in BHK than in HIT extracts. An 85-M<sub>r</sub> protein was seen only in BHK cells. The only band seen exclusively in HIT cells was 110 M<sub>r</sub> in size. Several other less abundant proteins of lower molecular weight from both HIT and BHK cells were crosslinked to the Pal-element sequence. However, these DNA-protein complexes were only minimally affected by the oligonucleotide competitor, indicating that the proteins were non-specific.

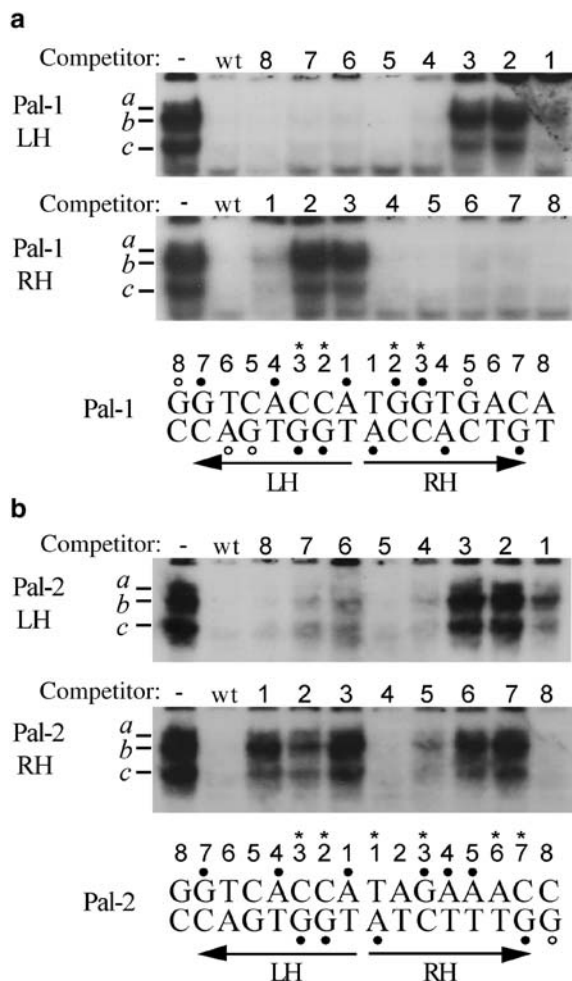


**Fig. 1.** Methylation interference contact point analysis of Pal-1 and Pal-2 binding proteins in HIT M2.2.2 and BHK cells. **a.** Probes containing Pal-1 sequences were incubated with either HIT (H)- or BHK (B)-cell nuclear extracts and the DNA-protein complexes were visualised by EMSA. **b.** The major complex from HIT cells (*b*) and BHK cells (*a*) were analysed by methylation interference analysis. The labelled top or bottom strand was complexed with the indicated nuclear proteins, cleaved and separated by denaturing PAGE. Interfered residues in the free (F) and bound (B) EMSA complexes are indicated by a horizontal line. \* An adenine contacted in HIT cells but not in BHK cells. **c.** Complexes resolved using the Pal-2 probe with nuclear extracts from HIT cells and BHK cells in the EMSA. **d.** Alignment of Pal-1 and Pal-2 regions with a summary of the contact points obtained (closed circles). Open circles indicate weaker contact points at specific purines. The arrows indicate the perfect inverted repeat in the Pal-1 region. The previously described Pal motif is in the 9-bp boxed region [4]

To determine whether the complexes observed by EMSA of the HIT nuclear proteins contained different ratios of the component proteins, a Pal-1 probe was crosslinked with proteins in a HIT nuclear extract and the proteins were resolved by EMSA. The individual complexes were excised and then separated by SDS gel electrophoresis. As shown in Fig. 2b, the *a* complex



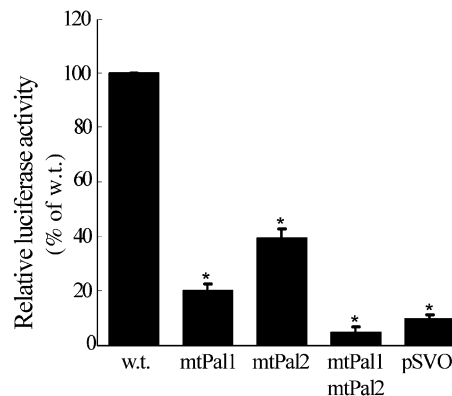
**Fig. 2.** UV-crosslinking of different complexes from HIT M2.2.2 and BHK cells to the Pal element. **a.** The BrdU-incorporated Pal-1 probe was crosslinked with nuclear extracts from HIT and BHK proteins in the absence (-) or presence (+) of competitor and resolved by SDS-PAGE. The 150-M<sub>r</sub> and 90-M<sub>r</sub> proteins were not cell-specific, while the 110-M<sub>r</sub> protein was specific to HIT cells. **b.** Complexes corresponding to bands *a*, *b* and *c* in EMSAs of HIT cell extract and Pal-1 (Fig. 1a) were excised and separated by SDS-PAGE. Complex *a* contains one major protein of 150 M<sub>r</sub>, complex *b* contains three major proteins of 150, 110, and 90 M<sub>r</sub>, and complex *c* contains two major proteins of 110 and 90 M<sub>r</sub>.



**Fig. 3.** Gel-shift competition assay testing the effect of single bp mutations on the Pal-binding complex. EMSAs showing the effects of a 100-fold excess of single point mutant oligonucleotide competitors on complexes *a*, *b* and *c* from HIT cells using the Pal-1 probe (**a**) or Pal-2 probes (**b**). The 24-bp oligonucleotide competitors contained either A↔C or G↔T transversion mutations at each of the numbered positions in Pal-1 or Pal-2 left half (LH) or right half (RH). The lanes contain either no competitor (–), wild-type (wt) competitor or one of the eight different point mutant competitors in the LH or RH of the element. \* Nucleotides critical for binding

was found to consist almost entirely of the 150-M<sub>r</sub> protein. In contrast, the *b* complex consisted of all three proteins (150, 110, and 90 M<sub>r</sub>), while the *c* complex was made up of only the 110-M<sub>r</sub> and 90-M<sub>r</sub> proteins.

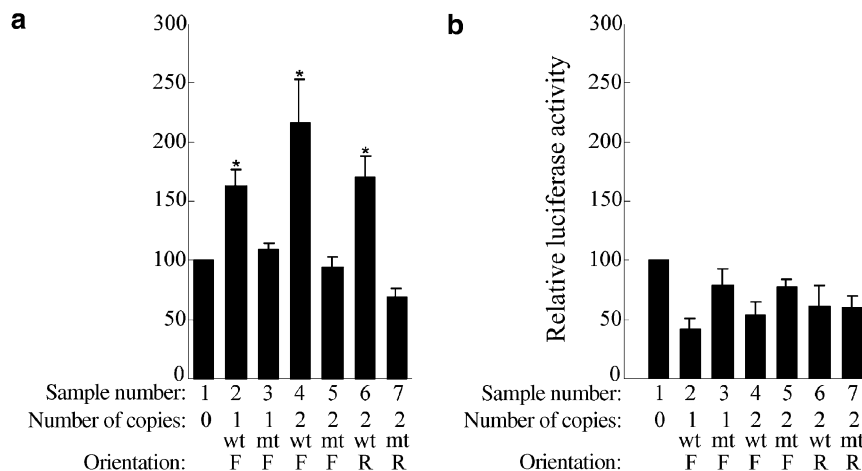
**Competition analysis of the major contact points.** To confirm that the nucleotides identified as contact points by methylation interference are also necessary for the formation of the Pal binding complex, a series of single point mutations were generated within the Pal-1 and Pal-2 elements, and the ability of the mutants to compete for protein binding was tested. Of the 16 mutants tested (Fig. 3), those with substitutions at positions 1, 4, 5, 6, 7 and 8 in the left half of the centre of symmetry were observed to partially or com-



**Fig. 4.** Functional analysis of mutations in the GK promoter in HIT M2.2.2 cells. The upstream GK promoter fragment fused to the luciferase reporter gene was mutated at a single nucleotide (C to A) at position 3 in the left half of either Pal-1 (mtPal1) or Pal-2 (mtPal2) or both. These constructs were transfected into HIT cells and luciferase activity was normalised to the activity of β-galactosidase. Transcriptional activity is shown relative to the wild type (w.t.; –402 to +14) promoter. The promoterless vector pSVO served as a negative control [4]. Values are the means ± SEM of at least four individual transfection experiments. \* *p* < 0.005 vs wild-type promoter

pletely compete for binding (Fig. 3a). In contrast, mutants with substitutions at positions 2 and 3 in the left half did not compete, and thus are the most critical for protein binding. In every instance where binding was competed, an equal reduction in all three complexes (*a*, *b* and *c*) was observed. The same analysis was applied to the right half of the Pal-1 element (Fig. 3a, lower panel). Essentially the same results were obtained with the Pal-1 right-half mutants: nucleotides at positions 2 and 3 in the right half of dyad symmetry were the most important for formation of a DNA–protein complex. The essential role of nucleotides at positions 2 and 3 (in the left half) for protein binding in HIT or BHK extracts was confirmed by generating DNA probes with mutations at these two positions. Use of these probes in EMSAs did not generate any DNA–protein complexes (data not shown). To assess the effects of contact points of the Pal-2 element on protein binding a similar analysis to that applied to Pal-1 was used. Mutants at positions 2 and 3 (in the left half) were the most critical for binding, as mutations at these sites prevented the competitor from binding to the Pal-2 probe (Fig. 3b, top). In contrast, the competition pattern in the right half of the Pal-2 element differed from the right half of Pal-1. A mutation at position 1 (in the right half) prevented competition with Pal-2, as did mutations at positions 6 and 7, suggesting that these nucleotides are important for protein binding.

**Effects of point mutations on βGK gene transcription.** Having established that a mutation at position 2 or 3 in the left half of Pal-1 or Pal-2 interferes with protein



**Fig. 5.** Effects of multimerised Pal elements on gene expression from a heterologous promoter. Tandem copies of the Pal-1 binding element (wt) in single or double copy and elements containing point mutations (mt) at position 3, previously shown to be critical for binding, were tested in HIT cells (a) and BHK cells (b). The Pal-1 element was cloned in the forward (F) or reverse (R) orientation. In HIT cells, the Pal sequences augmented transcription from the SV40 minimal promoter (lanes 2, 4 and 6) and the effect was abrogated in the mutant elements (lanes 3, 5 and 7). No augmentation from the minimal element was observed in BHK cells. The data are represented as percent luciferase activity from the SV40 minimal promoter vector (lane 1, no insert) normalised to the activity of  $\beta$ -galactosidase and are the means  $\pm$  SEM of three to nine independent experiments. \*  $p < 0.05$  vs SV40 minimal promoter construct

binding, and that these bases are fully conserved in both elements, we next tested the functional role of the elements on  $\beta$ GK promoter activity in HIT and BHK cells. A single mutation was introduced at position 3 in the left half of Pal-1 (mtPal1) and Pal-2 (mtPal2). As shown in Figure 4, the mutation in Pal-1 reduced transcription by 80% in HIT cells. The same mutation in the Pal-2 element reduced transcription by 61%. Together, the elements reduced transcription to a level similar to that observed for the promoterless vector. These results clearly demonstrate the essential role that both Pal elements play in the transcription of the  $\beta$ GK promoter in insulinoma-derived cells and probably pancreatic beta cells. Since protein complexes from BHK and other cells bind to the Pal elements in vitro, we investigated whether the Pal elements might function to repress transcription by analysing mtPal1 and mtPal2 in BHK cells. The transcriptional activity of a  $\beta$ GK promoter-reporter construct was lower in BHK cells than that of a promoterless vector. Moreover, the mutants did not have a de-repressive effect on the promoter (data not shown). Thus, these results do not support the notion that the Pal elements function as transcriptional repressors in cell types that do not express GK.

#### Role of single elements as transcriptional enhancers.

To determine whether the Pal elements exhibit enhancer activity when placed upstream of a minimal promoter, several new constructs containing multimers of the Pal-1 element were made and tested in HIT and BHK cells. As shown in Figure 5, constructs containing multimerised elements in both forward and reverse orientations augment transcription by a factor of 1.5 to 2 in HIT cells. In contrast, nearly identical constructs containing multimers of the Pal-1 element mutated at position 3 showed no increase above that of the promoter-only construct. The same result was obtained by transfecting the enhancer constructs in STC-1 enteroendocrine cells (data not shown). In BHK cells, the minimal element constructs had a small inhibitory effect on the minimal promoter in the forward orientation. Taken together, these results indicate that the Pal element, when multimerised, exhibits enhancer-like behaviour in HIT cells but not in BHK cells.

## Discussion

Previous studies have established that the two Pal motifs in the upstream  $\beta$ GK promoter are necessary for efficient transcription [4, 5]. However, these studies did not reveal the structure of the elements that contain these motifs. By using methylation interference analysis, we have determined the sequence of two closely related elements in the  $\beta$ GK promoter, which are now referred to as the Pal-1 and Pal-2 elements. While the Pal-1 element exhibits perfect dyad symmetry (GTCACCA-TGGTGA), the Pal-2 element (GTCACCA-TAGAAAC) differs on the right of the centre of symmetry. A systematic analysis using point mutations (utilising a base transversion strategy, e.g. A to C or G to T) revealed that bases 2 and 3 to the left and right of the centre of symmetry appear to be the most important for protein binding. Interestingly, the right half of Pal-2 contains a natural purine substitution (G to A) at position 2 that does not appear to impair its ability to bind proteins.

It is known that a single base mutation at position 3 (in the left half) is sufficient to eliminate binding to either Pal-1 or Pal-2. We exploited this fact to assess the effect of the two elements on the function of the  $\beta$ GK promoter in HIT cells. Single mutations at position 3 (in the left half) in either Pal-1 or Pal-2 resulted in a large decrease in transcription, while the double mutation essentially left the entire promoter functionless. As we have previously suggested, both Pal-1 and Pal-2 are critical for the expression of  $\beta$ GK in insulinoma cells, as well as in other neuroendocrine cells [5]. A multimerised repeat containing three Pal-1 elements functioned as a weak enhancer in HIT cells and STC-1 cells, but not in BHK cells. Thus, Pal element function appears to be specific to neuroendocrine cells.

The fact that the multimerised Pal-1 element increased transcription by a factor of only 1.5 to 2 when placed upstream of a minimal promoter suggests that the effect of these elements by themselves is weak and may be augmented by proteins that bind to flanking sequences in the  $\beta$ GK promoter. Indeed, a previous mutational analysis utilising a 3-bp block clearly demonstrated that sequences immediately upstream and downstream of Pal-2 are important for transcription of the  $\beta$ GK promoter in insulinoma cells [5]. For example another *cis*-element that is bound by the transcription factor Pdx-1 [6] is located only 9 bp upstream of the Pal-2 element, and the E-box element is located ~50 bp upstream. Pdx-1 has been shown to bind to and transactivate the promoters of several other genes expressed in the pancreatic islet, including those of the insulin, somatostatin, and *GLUT2* genes [17, 18, 19, 20]. However, since  $\beta$ GK is expressed in other NE cells such as pituitary corticotropes, and since these cells do not express Pdx-1, it is likely that different protein–protein interactions are necessary for transcriptional activation in different neuroendocrine cell types. This would be consistent with differences in the  $\beta$ GK promoter regulatory elements utilised in different neuroendocrine cells (our unpublished observations).

There are two Pal elements in the  $\beta$ GK promoter and both are necessary for full promoter activity. The positioning of these elements within the  $\beta$ GK promoter relative to that of other promoter elements such as the BETA2 and Pdx-1 binding sites may be important for conferring full transcriptional activity. Based on the results we obtained by multimerising the Pal sequences in a synthetic construct, we speculate that the positioning of these factors enables the PEBP complex to interact with the basal transcription machinery and augment basal transcription originating from the SV40 minimal promoter. The weak transcriptional enhancement observed with the multimerised Pal-1 element may also be related to the presence or lack of certain cell-specific co-activators [21], although cloning of the PEBPs will be necessary to explore this possibility.

Further computer analysis of the Pal sequences was performed to try and identify similar sequences in other gene regulatory regions. Using the MatInspector [22], we searched the transcription factor database TRANSFAC (<http://www.gene-regulation.com>, accessed in February 2004), but failed to find anything matching Pal-1 or Pal-2. However, a BLAST search of GenBank (<http://www.ncbi.nlm.nih.gov/blast>, in May 2004) using the 14-bp sequences of Pal-1 and Pal-2 resulted in the identification of approximately 190 sequences with 100% identity to the 14-bp Pal-1 element. However, most of these identity matches corresponded to sequences in mouse and human bacterial artificial chromosome clones that have not yet been fully annotated. A computer analysis of some of these sequences has not revealed whether they exist within gene regulatory elements. Two additional sequences were identified in the promoter regions of the human integrin gene [23] and chick *MyoD* [24], prompting us to evaluate them further at the experimental level. DNA oligonucleotides containing these sequences were able to compete for binding of the PEBPs in a gel-shift assay (data not shown), indicating that they also bind the PEBPs. However, it is not currently known whether these sequences are functionally important in the context of the *MyoD* and integrin promoters.

The results of our UV-crosslinking studies indicate that homodimers and heterodimers are formed on both Pal elements. Indeed, the formation of different protein complexes provides a possible explanation for our consistent observation that these elements are functional in neuroendocrine cells but not in other cells, despite the fact that similar proteins are expressed in both cell types. The formation of various homodimers and heterodimers would also help to explain the formation of complexes that differ in both size and abundance in HIT cells compared with BHK cells. For instance, formation of a homodimer of the 110-M<sub>r</sub> protein appears to only be possible in neuroendocrine cells, as BHK cells lack a protein of this size. The fact that DNA–protein interactions occur in cells that do not express GK implies that certain combinations of PEBPs are incapable of augmenting gene transcription. We proposed that certain combinations of PEBPs might actually be able to repress gene transcription in non-endocrine cells that was brought about, for example, by the 85-M<sub>r</sub> complex observed in BHK cells. However, our experiments failed to provide direct support for active inhibition or repression, although at this time we cannot exclude the possibility of gene occupancy by the PEBPs in cells that do not express  $\beta$ GK. Given that the 110-M<sub>r</sub> protein was identified only in HIT cells, this protein may be necessary for the transcriptional activity of Pal-1 and Pal-2 in this cell type. Clearly, further understanding of the interactions between different PEBPs and their ability to confer transcriptional activation in some cell types but

not others, requires cloning of the gene(s) encoding these proteins.

The size of the Pal elements may provide clues as to the identity of the transcription factors that bind to these elements. For instance, a family of transcription factors related to the *v-maf* oncogene, members of the bZIP superfamily, also recognise a 14-bp element (known as the Maf-recognition element or MARE) that is palindromic in structure [25]. However, members of the large Maf transcription factor family, which were recently identified as regulators of insulin gene expression, are still considerably smaller than the PEBPs, being approximately 46 M<sub>r</sub> in size [26]. Nevertheless, the Maf transcription factors may provide a model for understanding how the Pal element's function is affected by the binding of different proteins. Other potential candidates include members of the nuclear receptor superfamily, which may bind large elements that have a palindromic structure. In this respect, the Pal elements most closely resemble a thyroid-hormone-response element which is found in the promoter region of the  $\gamma$ F-crystallin gene and is an everted repeat [27, 28]. However, since members of this superfamily are typically less than 50 M<sub>r</sub> in size, it seems unlikely that they bind to the Pal elements. Moreover, we evaluated whether certain nuclear receptors bound to the Pal elements by using a DR1, DR3 and IR-1 binding site [29, 30] as competitors and were unable to compete the PEBPs in EMSAs (data not shown).

The sizes of the PEBPs as determined by UV-crosslinking analysis may provide additional information for the determination of the identity of the PEBPs. A search of the TRANSFAC database [22] for proteins with a molecular weight of 100, 110 and 150 M<sub>r</sub> revealed a number of known transcription factors with sizes similar to those of the PEBPs. Fifteen transcription factors were identified with molecular weights between 145 and 155 M<sub>r</sub> and 104 factors were identified with sizes between 95 and 115. Interestingly, there are two transcription factors, Friend of GATA-1 (FOG) which has both 106 and 150 M<sub>r</sub> isoforms [31] and the dioxin receptor which has 100 and 150 M<sub>r</sub> isoforms [32]. However, based on the sequences of the binding sites of these factors, it is unlikely that either of these are the PEBPs.

*Concluding comments.* In this study we determined the structure of two closely related elements in the  $\beta$ GK promoter. These elements, which are now termed the Pal elements, enhance transcription in neuroendocrine cells when multimerised. Given the importance of the Pal elements in the function of the  $\beta$ GK promoter, future experiments can now be directed at cloning the genes encoding the PEBPs and determining how these proteins regulate  $\beta$ GK gene expression in a restricted set of neuroendocrine cells. Furthermore, an understanding of all of the *cis*- and *trans*-

elements that regulate  $\beta$ GK gene expression is of fundamental importance, especially given the observation that  $\beta$ GK promoter variations, such as the -30 promoter polymorphism, are associated with Type 2 diabetes in humans [33].

*Acknowledgements.* This study was supported by grants from the Public Health Service (DK42502 and DK42612) to M. A. Magnuson. J. M. Moates was supported by a Research Career Development Award and a MERIT Review Grant from the Veterans Administration. There are no conflicts of interest in any of the presented work.

## References

1. Jetton TL, Liang Y, Pettepher CC et al. (1994) Analysis of upstream glucokinase promoter activity in transgenic mice and identification of glucokinase in rare neuroendocrine cells in the brain and gut. *J Biol Chem* 269:3641–3654
2. Yang XJ, Kow LM, Funabashi T, Mobbs CV (1999) Hypothalamic glucose sensor: similarities to and differences from pancreatic beta-cell mechanisms. *Diabetes* 48:1763–1772
3. Cheung AT, Dayanandan B, Lewis JT et al. (2000) Glucose-dependent insulin release from genetically engineered K cells. *Science* 290:1959–1962
4. Shelton KD, Franklin A, Khoor A, Beechem J, Magnuson MA (1992) Multiple elements in the upstream glucokinase promoter are necessary for transcription in insulinoma cells. *Mol Cell Biol* 12:4578–4589
5. Moates JM, Shelton KD, Magnuson MA (1996) Characterization of the Pal motifs in the upstream glucokinase promoter: Binding of a cell type-specific protein complex correlates with transcriptional activation. *Mol Endocrinol* 10:723–731
6. Watada H, Kajimoto Y, Umayahara Y et al. (1996) The human glucokinase gene b-cell-type promoter: An essential role of insulin promoter factor I/PDX-1 in its activation in HIT-T15 cells. *Diabetes* 45:1478–1488
7. Guz Y, Montminy MR, Stein R et al. (1995) Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. *Development* 121:11–18
8. Moates JM, Nanda S, Cissell MA, Tsai MJ, Stein R (2003) BETA2 activates transcription from the upstream glucokinase gene promoter in islet beta-cells and gut endocrine cells. *Diabetes* 52:403–408
9. Dumonteil E, Laser B, Constant I, Philippe J (1998) Differential regulation of the glucagon and insulin I gene promoters by the basic helix-loop-helix transcription factors E47 and BETA2. *J Biol Chem* 273:19945–19954
10. Poulin G, Turgeon B, Drouin J (1997) NeuroD1/beta2 contributes to cell-specific transcription of the proopiomelanocortin gene. *Mol Cell Biol* 17:6673–6682
11. Mutoh H, Fung BP, Naya FJ, Tsai MJ, Nishitani J, Leiter AB (1997) The basic helix-loop-helix transcription factor BETA2/NeuroD is expressed in mammalian enteroendocrine cells and activates secretin gene expression. *Proc Natl Acad Sci USA* 94:3560–3564
12. Liu M, Pereira FA, Price SD et al. (2000) Essential role of BETA2/NeuroD1 in development of the vestibular and auditory systems. *Genes Dev* 14:2839–2854



13. DeWet JR, Wood KV, DeLuca M, Helinski DR, Subramani S (1987) Firefly luciferase gene: structure and expression in mammalian cells. *Mol Cell Biol*:725–737
14. Shapiro DJ, Sharp PA, Wahli WW, Keller MJ (1988) A high-efficiency HeLa cell nuclear transcription extract. *DNA* 7:47–55
15. Cao J, Revzin A (1993) Interference and missing contact footprinting. Academic Press, San Diego
16. Chodosh LA, Carthew RW, Sharp PA (1986) A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol Cell Biol* 6:4723–4733
17. Ohlsson H, Karlsson K, Edlund T (1993) IPF-1, a homeodomain-containing transactivator of the insulin gene. *EMBO J* 12:4251–4259
18. Leonard J, Peers B, Johnson T, Ferreri K, Lee S, Montminy MR (1993) Characterization of somatostatin transactivating factor-1, a novel homeobox factor that stimulates somatostatin expression in pancreatic islet cells. *Mol Endocrinol* 7:1275–1283
19. Miller CP, McGehee RE Jr, Habener JF (1994) IDX-1: a new homeodomain transcription factor expressed in rat pancreatic islets and duodenum that transactivates the somatostatin gene. *EMBO J* 13:1145–1156
20. Waeber G, Thompson N, Nicod P, Bonny C (1996) Transcriptional activation of the GLUT2 gene by the IPF-1/STF-1/IDX-1 homeobox factor. *Mol Endocrinol* 10:1327–1333
21. McKenna NJ, Lanz RB, O'Malley BW (1999) Nuclear receptor coregulators: cellular and molecular biology. *Endocr Rev* 20:321–344
22. Heinemeyer T, Chen X, Karas H et al. (1999) Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. *Nucleic Acids Res* 27:318–322
23. Zutter MM, Painter AA, Staatz WD, Tsung YL (1995) Regulation of alpha 2 integrin gene expression in cells with megakaryocytic features: a common theme of three necessary elements. *Blood* 86:3006–3014
24. Dechesne CA, Wei Q, Eldridge J et al. (1994) E-box- and MEF-2-independent muscle-specific expression, positive autoregulation, and cross-activation of the chicken MyoD (CMD1) promoter reveal an indirect regulatory pathway. *Mol Cell Biol* 14:5474–5486
25. Blank V, Andrews NC (1997) The Maf transcription factors: regulators of differentiation. *Trends Biochem Sci* 22:437–441
26. Matsuoka TA, Zhao L, Artner I et al. (2003) Members of the large Maf transcription family regulate insulin gene transcription in islet beta cells. *Mol Cell Biol* 23:6049–6062
27. Tini M, Otulakowski G, Breitman M, Tsui L, Giguere V (1993) An everted repeat mediates retinoic acid induction of gF-crystallin gene: evidence of a direct role for retinoids in lens development. *Genes Dev* 7:295–307
28. Tini M, Tsui L, Guiguere V (1994) Heterodimeric interaction of the retinoic acid and thyroid hormone receptors in transcriptional regulation on the gF-crystallin everted retinoic acid response element. *Mol Endocrinol* 8:1494–1506
29. Naar A, Boutin J, Lipkin S et al. (1991) The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* 65:1267–1279
30. Umesono K, Murakami KK, Thompson CC, Evans RM (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 65:1255–1266
31. Tsang AP, Visvader JE, Turner CA et al. (1997) FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90:109–119
32. Mason GG, Witte AM, Whitelaw ML et al. (1994) Purification of the DNA binding form of dioxin receptor. Role of the Arnt cofactor in regulation of dioxin receptor function. *J Biol Chem* 269:4438–4449
33. Stone LM, Kahn SE, Fujimoto WY, Deeb SS, Porte D Jr (1996) A variation at position -30 of the beta-cell glucokinase gene promoter is associated with reduced beta-cell function in middle-aged Japanese-American men. *Diabetes* 45:422–428