

# Functional analysis of a newly identified TAAT-box of the rat insulin-II gene promoter

Ingo B. Leibiger, Thomas Schwarz, Barbara Leibiger, Reinhard Walther\*

*Institut für Biochemie, Klinikum, Medizinische Fakultät, Ernst-Moritz-Arndt Universität, Sauerbruchstraße, D-17489 Greifswald, Germany*

Received 21 January 1995; revised version received 28 February 1995

**Abstract** Transcriptional regulation of insulin gene expression is achieved by an interplay of tissue-specific and ubiquitous *cis*- and *trans*-acting elements. E-box like motifs and TAAT-motifs were shown to play a crucial role in initiating insulin gene transcription. Studying the AT-rich region of the rat insulin-II promoter between nucleotides –212 and –196, we observed a base difference at –211, an adenosine instead of a cytidine, compared to the previously reported sequence (EMBL Accession No. J00748). Sequence analysis of promoter fragments from different rat strains showed that adenosine at position –211 represents the wild type (EMBL Accession No. X82162). This base exchange leads to the formation of an additional TAAT-motif, i.e. TAAT3, at the complementary DNA strand directly upstream of the previously studied TAAT2 motif, formerly named CT-2. Here we show that the newly identified motif TAAT3 is involved in (i) transcriptional control *in vivo*, (ii) *in vitro* DNA/protein interactions, and that (iii) TAAT1, TAAT2 and TAAT3 are binding sites for the homeodomain-containing factor IPF-1.

**Key words:** Rat insulin II gene; Gene expression; Transcriptional control; Insulinoma; DNA–protein interaction; Promoter

## 1. Introduction

The expression of the insulin gene in adult organisms is restricted to pancreatic  $\beta$ -cells. Multiple elements exhibiting both positive and negative effects contribute to its transcriptional control (for review see [1–3]). At least two known *cis*-elements seem to play a crucial role in both the tissue-specific and the glucose-responsive regulation of insulin gene transcription: E-box like sequences with a consensus sequence GCCANNTG, and TAAT-boxes. The E-box like motifs such as IEB2/Far- and IEB1/Nir-boxes [4,5] in the rat insulin-I promoter, ICE [6] and part of RIPE3 [7] in the rat insulin-II promoter, and the GC-I motif in the human insulin promoter [8] are binding sites for factors of the basic helix-loop-helix family. Whereas this class of transcriptional control elements is relatively well studied in insulin gene regulation, the investigation of the TAAT-motifs, representing binding sites for homeodomain-containing factors, was so far mainly restricted to the rat insulin-I gene [9,10] and the human insulin gene [8,11–15].

In a previous report we showed the impact of two TAAT-motifs between nucleotides –82/–76 and –202/–196 (CT-1 and CT-2, respectively) of the rat insulin II promoter in transcription initiation in insulin producing cells [16]. Both elements

were mutationally sensitive in transient expression studies and showed a similar retardation pattern in electrophoretic mobility shift assays. In a very recent study Peshavaria et al. [17] confirmed the data for the distal region between –201 and –196 bp and showed that it represents a binding site for the homeodomain containing factor STF-1/IPF-1/IDX-1. Furthermore, studying the region containing the CT-2 element we observed a difference in the nucleotide sequence at promoter position –211: an adenosine instead of the so far described cytidine. This base exchange would lead to the formation of an additional TAAT-motif at the lower DNA-strand directly upstream of the upper-strand distal TAAT-motif (CT-2). In this study we demonstrate that the newly identified TAAT-motif, called here TAAT3, represents the ‘wild type’ situation of this promoter region and, further, that TAAT3 is involved in (i) transcriptional control *in vivo*, (ii) *in vitro* DNA–protein interactions, as demonstrated by DNase I footprinting analysis and methylation interference assays, and (iii) represents a binding site for the homeodomain-containing factor IPF-1.

## 2. Material and methods

### 2.1. DNA-preparation and sequencing

Genomic DNA from Wistar and Sprague–Dawley rats was prepared as described in [18]. DNA from BB- and DA-rats were gifts from Dr. I. Klöting (Institut für Diabetes, EMAU Greifswald, Karlsburg). To amplify DNA-fragments of the rat insulin-II promoter using PCR we used as the upstream primer 5'-CTCTCTGAGACAATGTC-3' and as the downstream primer 5'-GTTACTGAACCCCACT-3', and *Ultma*-polymerase (Perkin Elmer) or Taq-Polymerase (Amersham) as thermostable DNA-polymerases. Generated PCR products were subcloned directly into the T-tailed vector pT7Blue (Novagen) and were verified by either cycle sequencing using the *Dye Termination Technique* and the 373A-sequencer (Applied Biosystems) or the T7-DNA-polymerase sequencing kit (Pharmacia). The nucleotide sequence data reported in this paper was deposited under accession number X82162 in EMBL/GenBank/DDJB.

### 2.2. Plasmid construction

Both plasmids p211A and p211C contain the rat insulin-II gene promoter (–218bp/+129bp) fused to the reporter gene CAT, but differ in the promoter sequence at position –211 in containing an adenosine (p211A) or a cytidine (p211C). p211A is identical with pBRINSII-218CAT [16]. p211C was generated on the basis of pBRINSII-218CAT by site-directed mutagenesis using the MutaGene *In Vitro* Mutagenesis Kit (Bio-Rad) and with the mutagenic primer 5'-CTTAGGGTAATAGAGTCTTAAGAGGGGCT-3'. After performing mutagenesis the promoter fragment was subcloned into the promoterless vector pBSV0CAT [19] and the sequence was verified.

### 2.3. Cell culture and transfection

Insulin-producing cells, HIT M2.2.2, were grown in RPMI 1640 supplemented with 40 units/ml penicillin, 40  $\mu$ g/ml streptomycin, 2 mM glutamine, and 10% fetal calf serum. Cells were transfected using the calcium phosphate coprecipitation technique as described in [20] using 20  $\mu$ g CsCl-purified CAT expression vector and 5  $\mu$ g pRSV $\beta$ gal [21] per 100 mm dish. Cells were harvested 48 h after transfection, and cell

\*Corresponding author. Fax: (49) 3834 88 3328; e-mail: leibiger@biochemie.medizin.uni-greifswald.d400.de



<b>R2AT211A</b>	AGCCCCTATTAAGACTCTAATTACCCTAAG TCGGGGATAATTCTGAGATTAATGGGATTC	-218/-189
<b>R2AT211C</b>	AGCCCCTcTTAAGACTCTAATTACCCTAAG TCGGGGAgAATTCTGAGATTAATGGGATTC	-218/-189
<b>R2TAAT3</b>	AGCCCCTATTAAGACgacgAcTACCCTAAG TCGGGGATAATTCTGctgcTgATGGGATTC	-218/-189
<b>R2TAAT2</b>	AGCCCCTgcTgAGACTCTAATTACCCTAAG TCGGGGAcgAcTCTGAGATTAATGGGATTC	-218/-189
<b>R2TAAT1</b>	TGCTGATCCACCCTTAATGGGACAAACAGC ACGACTAGGTGGGAATTACCCTGTTTGTGG	-94/-65
<b>R1FLAT</b>	CCTTGTTAATAATCTAATTACCCTAGGTCT GGAACAATTATTAGATTAATGGGATCCAGA	-227/-198
<b>R1TAAT1</b>	CCTCCTAGAGCCCTTAATGGGCCAAACGGC GGAGGATCTCGGGAAATTACCCGTTTGGCC	-94/-65

Fig. 3. Position and sequence of oligonucleotide probes used for the analysis of protein binding in EMSAs and competition studies. Oligonucleotides contain TAAT-motifs of the rat insulin-I promoter (R1TAAT1, R1FLAT) and normal (R2TAAT1, R2AT211A) and mutated (R2AT211C, R2TAAT2, R2TAAT3) regions of the rat insulin-II promoter. Positions of mutations within mutant oligonucleotides are shown as lower-case letters.

insulin-II promoter between nucleotides  $-216/-187$ , i.e. R2AT211A, led to the formation of three retarded complexes A, B, and C (Fig. 4A, lane 1). Surprisingly, the oligonucleotide containing the originally described sequence composition, i.e. R2AT211C, generated the same retardation pattern as seen from R2AT211A (Fig. 4A, lane 6). However, while changing the incubation conditions, especially the concentration of KCl, we observed a marked difference in the retardation patterns of both probes: whereas increasing amounts of KCl from 16 mM to 200 mM led to a stronger formation of complex B for the wild-type variant R2AT211A (Fig. 4A, lanes 1–5) the same conditions caused at best a weak increase in the formation of complex B for R2AT211C (Fig. 4A, lanes 6–10). To study the protein-binding properties of the newly identified TAAT-motif on its own, we synthesized an oligonucleotide where the adjacent upper-strand TAAT-motif was mutated, i.e. R2TAAT3 (Fig. 3). In EMSAs R2TAAT3 formed only a single retarded complex (Fig. 4B, lane 4), very similar to complex B of R2AT211A (Fig. 4B, lane 1). The formation of this complex was abolished when using oligonucleotides containing the TAAT2-motif (R2TAAT2) or the TAAT1-motif (R2TAAT1) as competitors (Fig. 4B, lanes 6 and 7, respectively). Furthermore, R2TAAT3 as a competitor was able to reduce the formation of all complexes A, B, and C (Fig. 4B, lane 3).

#### 3.4. The AT-rich region between $-212$ and $-196$ bp of the rat insulin-II promoter contains two binding sites for IPF-1

IPF-1 was shown to be the favored *trans*-acting factor binding to the proximal TAAT-motif P1 of the rat insulin-I promoter but not to the distal E2 element (i.e. the FLAT-element [10]), where IEF2 binds when using nuclear extracts of the insulin-producing cell line  $\beta$ TC-1 [9].

To study whether the AT-rich region at  $-212/-196$  bp of the

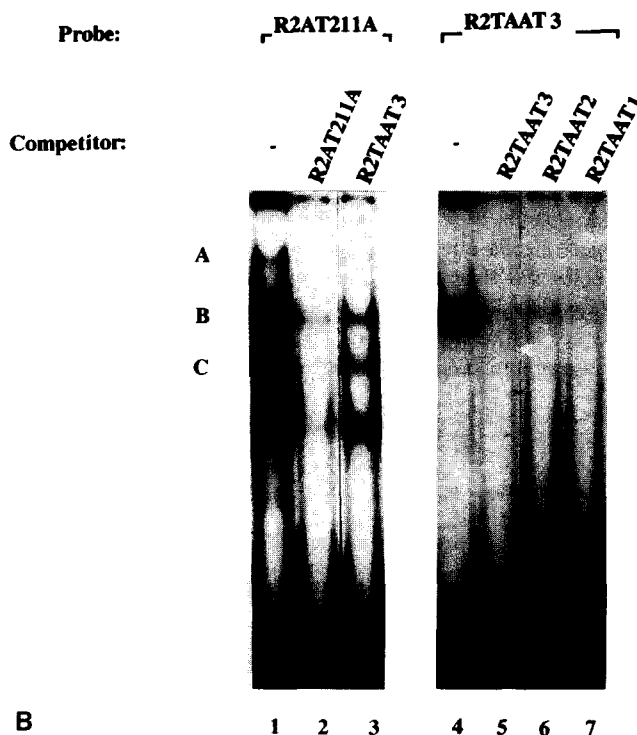
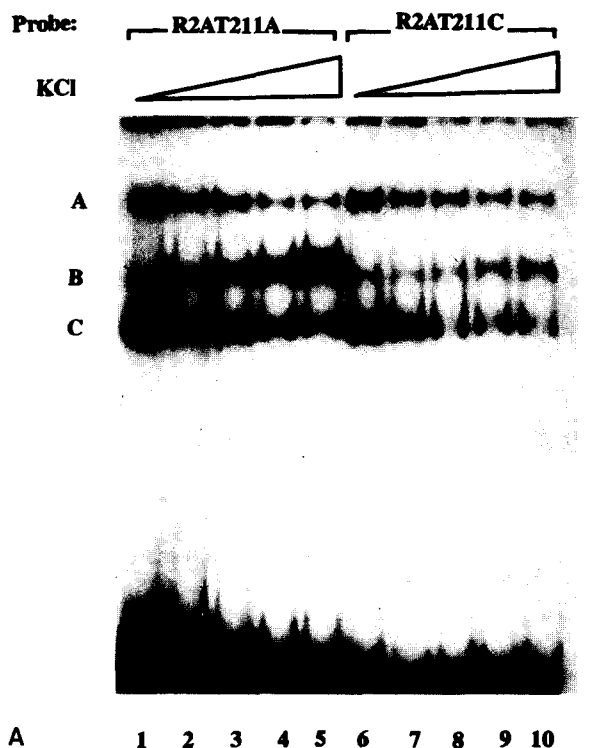
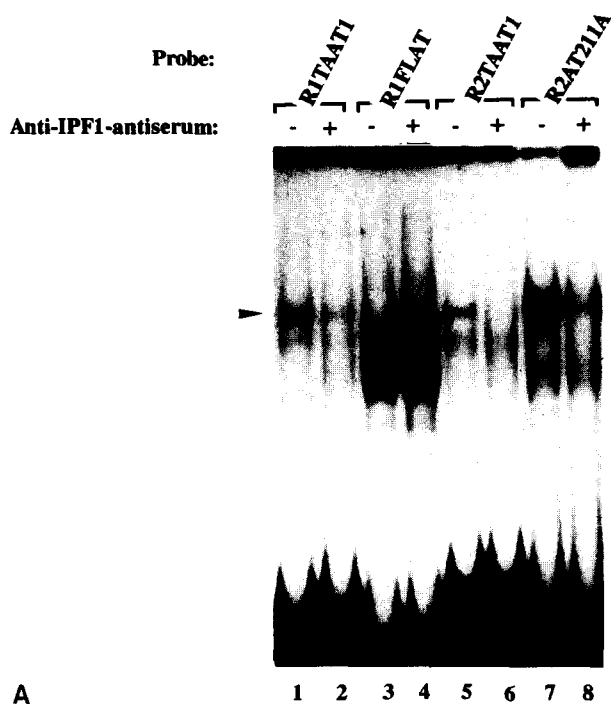


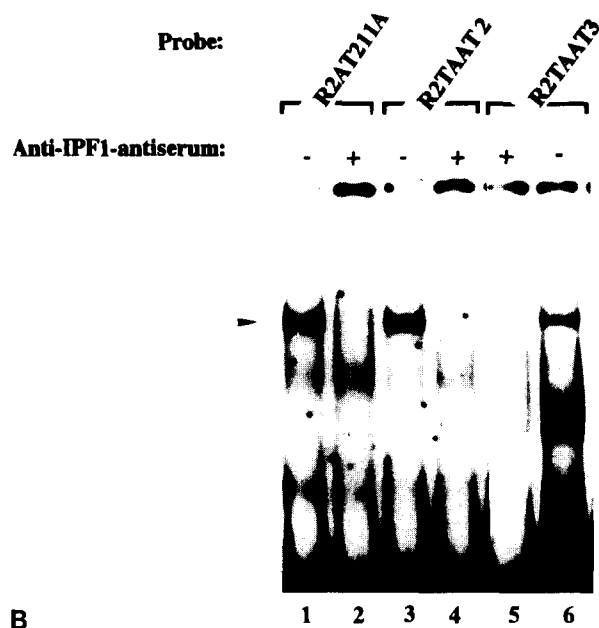
Fig. 4. Binding of HIT M 2.2.2 nuclear proteins to oligonucleotides R2AT211A, R2AT211C, and R2TAAT3. (A) Labeled oligonucleotides R2AT211A (lanes 1–5) and R2AT211C (lanes 6–10) were incubated with nuclear proteins in the presence of increasing amounts of KCl: 16 mM (lanes 1 and 6), 50 mM (lanes 2 and 7), 100 mM (lanes 3 and 8), 150 mM (lanes 4 and 9), and 200 mM (lanes 5 and 10), respectively. (B) Labeled oligonucleotides R2AT211A (lanes 1–3) and R2TAAT3 (lanes 4–7) were incubated in the presence of 150 mM KCl. The indicated unlabeled competitor oligonucleotides were used in approximately 100-fold molar excess.

rat insulin-II promoter is a target of IPF-1, we performed electrophoretic mobility shift assays using nuclear extracts pre-



A

1 2 3 4 5 6 7 8



B

1 2 3 4 5 6

Fig. 5. Binding of  $\beta$ TC-3 nuclear proteins to oligonucleotide probes corresponding to TAAT-motifs of rat insulin-I and -II promoters (see Fig. 3). Protein binding to the indicated labeled probes was performed with (+) or without (-) a pre-incubation of the nuclear extracts with anti-IPF-1 antiserum for 10 min at room temperature. The arrowheads indicate the position of IPF-1-containing complexes. (A) Binding of treated and untreated nuclear proteins to oligonucleotides containing the proximal or distal TAAT-motifs of rat insulin-I promoter (R1TAAT1 and R1FLAT, respectively), and rat insulin-II promoter (R2TAAT1 and R2AT211A, respectively). (B) Binding of treated and untreated nuclear proteins to oligos containing the intact distal AT-rich region with the combination of both TAAT2 and TAAT3 motifs (R2AT211A), or to mutant oligos containing either the TAAT2-motif (R2TAAT2) or TAAT3-motif (R2TAAT3) alone.

pared from  $\beta$ TC-3, a cell line very similar to  $\beta$ TC-1 [29]. As a positive control we used an oligonucleotide containing the proximal TAAT-motif (-80/-74 bp) of the rat insulin-I promoter (R1TAAT1), and as a negative control the FLAT-element (R1FLAT), i.e. the rat insulin-I promoter AT-rich region between -222/-208 bp. The EMSAs revealed similar retardation pattern for oligonucleotides containing the rat insulin-I IPF-1 binding site (R1TAAT1), the proximal TAAT-motif (R2TAAT1) and the AT-rich region between -212 and -196 bp (R2AT211A) of the rat insulin-II promoter (Fig. 5A, lanes 1, 5, 7, respectively). Addition of antibodies raised against the DNA-recognition part of IPF-1 lead to a loss in DNA-binding to both rat insulin-II AT-rich regions (Fig. 5A, lanes 6, 8). On the contrary, in the rat insulin-I promoter only the proximal TAAT-motif showed a loss in binding of IPF-1 when using the antibody (Fig. 5A, lane 2). When the IPF-1 binding properties of the two TAAT-motifs of the rat insulin-II AT-rich region were studied separately, both TAAT-motifs were able to bind protein complexes, which disappeared upon inclusion of antibodies raised against IPF-1 (Fig. 5B, lanes 4, 5). As expected, the FLAT-element of the rat insulin-I promoter formed retarded complexes different from those seen with the equivalent AT-rich region of the rat insulin-II promoter (Fig. 5A, lane 3), and these complexes were insensitive to treatment with anti-IPF-1 antibodies (Fig. 5A, lane 4).

#### 4. Discussion

Previous studies on insulin gene regulation by different groups have confirmed the involvement of TAAT-motifs in transcriptional control (8–17,30–32). TAAT-motifs were shown to be the binding core for homeodomain-containing transcription factors [33]. Several *trans*-acting factors were reported to bind to insulin TAAT-motifs: IUF-1 [11,12,14], *isl-1* [31], IPF-1 [26] and IEF-2 [9], *lmx-1* and *cdx-3* [32], STF-1/IPF-1 [15], STF-1/IPF-1/IDX-1 [17]. In a previous report we demonstrated that TAAT-motifs of the rat insulin-II promoter between nucleotides -83/-76 and -204/-197 are mutationally sensitive and contribute to transcriptional control [16]. Furthermore, we found a difference in the nucleotide composition within the AT-rich region at -212/-196 at position -211 compared to the originally described sequence [27], i.e. an adenosine instead of a cytidine. One consequence of that sequence difference is the formation of an additional TAAT-motif on the complementary strand directly upstream of the well studied TAAT2-motif [16]. Here we present evidence that this sequence represents the 'wild type' situation for the rat insulin-II promoter for at least four different rat strains: Wistar, Sprague-Dawley, DA, and BB-rats. As demonstrated by in vitro DNA-protein interaction analysis this region is protected from DNase I digestion and is also involved in methylation interference (Fig. 2). Although the observed sequence difference at position -211 did not lead to a changed retardation pattern in EMSA at 16 mM KCl, the formation of these complexes is clearly different under increasing concentrations of KCl (Fig. 4A). This might be the reason for the more than 2.3-fold difference in CAT expression using constructs with both sequence variants (Fig. 1A). The differences in retardation patterns and binding affinities of the same oligonucleotide probes in figures 4 and 5 is caused by the use of nuclear extracts of different species, hamster insulinoma (Fig. 4) and mouse insulinoma (Fig. 5). Con-

cerning the involvement of the TAAT2- motif in transcriptional control, our previous observation [16] was confirmed by Peshavaria et al. [17] showing furthermore that this motif is a binding site for a homeodomain-containing protein very similar to STF-1/IPF-1/IDX-1 and related to the endoderm-specific *Xenopus laevis* homeodomain protein XIHbox8. Since the insulin promoter factor 1, i.e. IPF-1, was demonstrated to be identical to a part of XIHbox8 [26] and to be the favored binding factor for the proximal rat insulin-I TAAT-motif, i.e. to the P1-region [9], we investigated the binding of IPF-1 to the AT-rich region of the rat insulin-II promoter located between nucleotides –212 and –196. Using nuclear extracts prepared from  $\beta$ TC-3 as well as antibodies raised against IPF-1 in electrophoretic mobility shift assays we demonstrate that not only the previously reported TAAT-motifs TAAT1 and TAAT2, but also the newly identified motif TAAT3, is a binding target of IPF-1. Additionally our studies confirm the findings by Ohlsson et al. [9] showing that the two TAAT-containing regions within the rat insulin-I promoter, i.e. P1 and E2 (FLAT), represent binding sites for different *trans*-acting factors. Our observations are also in agreement with those of German et al. [10] that different *trans*-acting factors are involved in complex-formation of the AT-rich regions of the rat insulin-I and -II promoters. The fact, that TAAT2/TAAT3 is a binding site for IPF-1, and the absence of a correct Far-element (GCCATCTG) in the correct juxtaposition to the AT-rich region of the rat insulin-II promoter, may contribute to the marked difference observed in the transcriptional control of both rat insulin genes, i.e. the promoter region of the rat insulin-II gene similar to the rat insulin-I FF-minienhancer does not exhibit enhancer properties ([10], own observations, data not shown). On the other hand we found that IPF-1 binding is not restricted to the promoters of different insulin genes. IPF-1 is also a *trans*-acting factor binding to TAAT-motifs of further  $\beta$ -cell-active promoters, such as the rat glucokinase upstream promoter, the rat amylin gene promoter, and the human glucose transporter-2 gene promoter (I.B. Leibiger, unpublished data). These findings support our previous assumption, based on EMSA studies, that TAAT-motifs of the genes mentioned above may use similar, if not identical, *trans*-acting factors [25]. In this way, IPF-1, or related factors such as IUF-1 [11], STF-1 [34], or IDX-1 [35], may represent one of the central factors responsible for  $\beta$ -cell-active and islet-specific transcriptional control, and, furthermore, may be involved in the coordinated regulation of pancreatic  $\beta$ -cell expressed genes.

*Acknowledgements:* This work was supported in part by grants from Deutsche Forschungsgemeinschaft (Wa767/1-3 and Le804/1-1).

## References

- [1] Stein, R. (1993) Trends Endocrinol. Metab. 4, 96–101.
- [2] Shiran, R., Aronheim, A., Rosen, A., Park, C.W., Leshkowitz, D. and Walker, M.D. (1993) Biochem. Soc. Trans. 21, 150–154.

- [3] Docherty, K. and Clark, A.R. (1994) FASEB J. 8, 20–27.
- [4] Karlsson, O., Walker, M.D., Rutter, W.J. and Edlund, T. (1989) Mol. Cell. Biol. 9, 823–827.
- [5] Moss, L.G., Moss, J.B. and Rutter, W.J. (1988) Mol. Cell. Biol. 8, 2620–2627.
- [6] Whelan, J., Cordle, S.R., Henderson, E., Weil, P.A. and Stein, R. (1990) Mol. Cell. Biol. 10, 1564–1572.
- [7] Shieh, S.-Y. and Tsai, M.-J. (1991) J. Biol. Chem. 266, 16708–16714.
- [8] Boam, D.S.W., Clark, A.R. and Docherty, K. (1990) J. Biol. Chem. 265, 8285–8296.
- [9] Ohlsson, H., Thor, S. and Edlund, T. (1991) Mol. Endocrinol. 5, 897–904.
- [10] German, M.S., Moss, L.G., Wang, J. and Rutter, W.J. (1992) Mol. Cell. Biol. 12, 1777–1788.
- [11] Boam, D.S.W. and Docherty, K. (1989) Biochem. J. 264, 233–239.
- [12] Scott, V., Clark, A.R., Hutton, J.C. and Docherty, K. (1991) FEBS Lett. 290, 27–30.
- [13] Clark, A.R., Petersen, H.V., Read, M.L., Scott, V., Michelsen, B.K. and Docherty, K. (1993) FEBS Lett. 329, 139–143.
- [14] MacFarlane, W.M., Read, M.L., Gilligan, M., Bujalska, I. and Docherty, K. (1994) Biochem. J. 303, 625–631.
- [15] Petersen, H.V., Serup, P., Leonard, J., Michelsen, B.K. and Madsen, O. (1994) Proc. Natl. Acad. Sci. USA 91, 10465–10469.
- [16] Rommel, C., Leibiger, I.B., Leibiger, B. and Walther, R. (1994) FEBS Lett. 345, 17–22.
- [17] Peshavaria, M., Gamer, L., Henderson, E., Teitelman, G., Wright, C.V.E. and Stein, R. (1994) Mol. Endocrinol. 8, 806–816.
- [18] Leibiger, B., Leibiger, I., Sarrach, D. and Zühlke, H. (1991) Biochem. Biophys. Res. Commun. 174, 1223–1231.
- [19] Clark, A.R., Boam, D.S. and Docherty, K. (1989) Nucleic Acids Res. 17, 10130.
- [20] Graham, F.L. and van der Eb, A.J. (1973) Virology 52, 456.
- [21] Edlund, T., Walker, M.D., Barr, P.J. and Rutter, W.J. (1985) Science 230, 912–916.
- [22] Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044.
- [23] Herbomel, P., Bourachot, B. and Yaniv, M. (1984) Cell 39, 653.
- [24] Schreiber, E., Matthias, P., Müller, M.M. and Schaffner, W. (1989) Nucleic Acids Res. 17, 6419.
- [25] Leibiger, B., Walther, R. and Leibiger, I.B. (1994) Biol. Chem. Hoppe-Seyler 375, 93–98.
- [26] Ohlsson, H., Karlsson, K. and Edlund, T. (1993) EMBO J. 12, 4251–4259.
- [27] Soares, M.B., Schon, E., Henderson, A., Karathanasis, S.K., Cate, R., Zitlin, S., Chirgwin, J. and Efstradiadis, A. (1985) Mol. Cell. Biol. 5, 2090–2103.
- [28] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463.
- [29] Efrat, S., Linde, S., Kofod, H., Spector, D., Delannoy, M., Grant, S., Hanahan, D. and Baekkeskov, S. (1988) Proc. Natl. Acad. Sci. USA 85, 9037–9041.
- [30] Crowe, D.T. and Tsai, M.-J. (1989) Mol. Cell. Biol. 9, 1784–1789.
- [31] Karlsson, O., Thor, S., Norberg, T., Ohlsson, H. and Edlund, T. (1990) Nature 344, 979–882.
- [32] German, M.S., Wang, J., Chadwick, R.B. and Rutter, W.J. (1992) Genes Dev. 6, 2165–2176.
- [33] Carton, K.M., Iler, N. and Abate, C. (1993) Mol. Cell. Biol. 13, 2354–65.
- [34] Leonard, J., Peers, B., Johnson, T., Ferrere, K., Lee, S. and Montminy, M.R. (1993) Mol. Endocrinol. 7, 1275–1283.
- [35] Miller, C.P., McGehee Jr., R.E. and Habener, J.F. (1994) EMBO J. 13, 1145–1156.