Cloning and DNA-binding properties of the rat pancreatic β-cell-specific factor Nkx6.1

Mette Christine Jørgensen[a], Helle Vestergård Petersen[a], Johan Ericson[b], Ole Dragsbæk Madsen[a], Palle Serup[a,*]

[a] *Department of Developmental Biology, Hagedorn Research Institute, Niels Steensensvej 6, DK-2820 Gentofte, Denmark*
[b] *Department of Cell and Molecular Biology, Karolinska Institute, Doktorsringen 2D, S-17177 Stockholm, Sweden*

Received 25 September 1999

Abstract The homeodomain (HD) protein Nkx6.1 is the most β-cell-specific transcription factor known in the pancreas and its function is critical for the formation of the insulin-producing β-cells. However, the target genes, DNA-binding site, and transcriptional properties of Nkx6.1 are unknown. Using in vitro binding site selection we have identified the DNA sequence of the Nkx6.1 binding site to be TTAATTG/A. A reporter plasmid containing four copies of this sequence is activated by an Nkx6.1HD/VP16 fusion construct. Full-length Nkx6.1 fails to activate this reporter plasmid in spite of robust interaction with the binding site in vitro. Stable expression of Nkx6.1 in the glucagon-producing α-cell-like MSL-G-AN cells induces expression of the endogenous insulin gene in a subset of the cell population. The expression of other known β-cell-specific factors such as Pax4, Pax6, Pdx1, GLUT2 and GLP1-R is unchanged by the introduction of Nkx6.1.

Key words: Nkx6.1; Pancreatic β-cell; Cloning; DNA-binding site

1. Introduction

The homeodomain protein Nkx6.1 was originally identified in a screen for new pancreatic islet cell homeobox genes by degenerate PCR using a HIT (hamster insulinoma) cDNA library [1]. During rat embryonic development, Nkx6.1 is detected in almost all of the nuclei of the epithelial cells in both of the pancreatic buds. Early endocrine cells are devoid of Nkx6.1 while later, mature insulin-producing β-cells display a strong nuclear Nkx6.1 immunoreactivity [2]. In the newborn rat pancreas, Nkx6.1 expression is restricted to the nuclei of the β-cells of the islets. Additionally, Nkx6.1 is expressed in the antral part of the stomach [3,4]. Furthermore, studies in mouse and chick embryos have shown that Nkx6.1 is expressed in the central nervous system (CNS) during development where it is found in the ventral part of the neural tube with a clear anterior boundary near the zona limitans intrathalamica [5]. This restricted expression pattern of Nkx6.1 suggests that Nkx6.1 is necessary for proper β-cell formation and CNS development. Nkx6.1 belongs to the NK family of homeodomain transcription factors even though the relationship to the majority of the family members is weak. The closest relative is Nkx6.2 (also called Gtx, glial- and testis-specific homeobox gene) [6], which shares an overall 67.5% sequence identity with Nkx6.1. The long alanine and serine stretches found in the N-terminal part of Nkx6.1 are absent in Nkx6.2, whereas the homeodomains are 95% identical (only three residues differ). The gene products of more thoroughly characterised NK family members appear to be involved in specification of cell fate and differentiation in specific tissues. This is exemplified by Nkx2.1 (also known as TTF1 and T/EBP), which is required for proper development of the thyroid, lungs, and the basal plate of the forebrain [7], and Nkx2.2 which is required for proper patterning of the neural tube [8] and for the differentiation of completely matured pancreatic β-cells [9].

Little is known about the control of Nkx6.1 expression. In mice where the Pdx1 gene was specifically inactivated in the β-cells, Nkx6.1 could not be detected in the pancreatic islets [10]. Similarly, Nkx2.2-deficient mice show a complete lack of Nkx6.1 in the islet β-cells, which also lack insulin expression [9]. In mature β-cells, Nkx6.1 expression seems to require the presence of both Pdx1 and Nkx2.2. However, nothing is yet known about target genes downstream of Nkx6.1.

To begin to address the function of Nkx6.1, we have isolated the rat Nkx6.1 cDNA sequence and determined the optimal DNA-binding site of the protein. A reporter plasmid with four copies of the Nkx6.1-binding site inserted in front of a TK promoter was activated by an Nkx6.1HD/VP16AD fusion protein. In an attempt to discover target genes for Nkx6.1, we have stably transfected MSL-G-AN glucagonoma cells with an Nkx6.1 cDNA and studied the expression of multiple β-cell-specific genes in the resulting clones.

2. Materials and methods

2.1. Screening of cDNA library

A 220 bp sequence fragment of Nkx6.1 was PCR amplified from rat islet cDNA using the hamster Nkx6.1 primers 5′-TCCCGGGAATTCATGGCCACCGCCAAGAAGAAGCAG-3′ (sense) and 5′-GCGGGCGGGCGGCCGCCGCTCAGGACGAGCCCTCGG-3′ (antisense) containing the restriction sites for *Eae*I and *Not*I, respectively. The fragment was cut with *Eae*I and *Not*I and inserted at the *Not*I site of pBluescript SK II$^+$ from Stratagene. The correct identity of the insert was verified by sequencing and used as the Nkx6.1 specific probe. It was ^{32}P-labeled with Klenow and random priming using the RPN 1601Y Multiprime DNA labeling system from Amersham. A high stringency colony hybridisation screening of 10^5 cfu was performed essentially as described in [11] on a newborn rat islet cDNA library [12].

2.2. Sequencing of rat Nkx6.1 cDNA

The 3.2 kb rat Nkx6.1 cDNA sequence was obtained by primer walking and was carried out doing cycle sequencing PCR using the fmol DNA Sequencing System from Promega or Thermo Sequenase from Amersham. The reactions were separated on 5% polyacrylamide

*Corresponding author. Fax: (45) 4443 8000.
E-mail: pas@hagedorn.dk

PII: S0014-5793(99)01436-2

sequencing gels with 8.3 M urea in 1×TBE. Both strands of the coding region were sequenced, whereas the untranslated regions (UTRs) were only sequenced in one direction.

2.3. In vitro translation

The 3211 bp Nkx6.1 cDNA sequence was inserted in the *Xba*I site of the pBK-RSV vector from Stratagene. In vitro transcription and translation of Nkx6.1 was carried out using the TNT Lysate Coupled Transcription/Translation System from Promega with the T3 RNA polymerase and in the presence of [^{35}S]methionine. The samples were analysed on SDS-PAGE and autoradiographed.

2.4. Expression of Nkx6.1 in COS7 cells

The 3211 bp Nkx6.1 cDNA sequence was inserted in the *Xba*I site of the pCMV4 vector [13]. 3.5×10^5 COS7 cells were mixed with 1 μg pCMV4-Nkx6.1 (3.2 kb) or 1 μg pCMV4 control vector and transfected by electroporation at 700 V for 90 μs with a 0.2 cm electrode gap using an Electro Square Porator T820 from BTX Electroporation Systems. The cells were allowed to recover for 24 h and then resuspended in 150 μl lysis buffer containing 20 mM imidazole-HCl pH 6.8, 100 mM KCl, 1 mM $MgCl_2$, 10 mM EGTA, 0.2% Triton X-100 and 1 mM AEBSF. Incubation for 5 min at room temperature and centrifugation for 5 min at $14\,000\times g$ gave a crude protein extract preparation of the transfected COS7 cells. Western blot analysis was performed as described [14].

2.5. GST-Nkx6.1HD fusion protein

The GST-Nkx6.1HD fusion protein was generated using the GST Gene Fusion System from Pharmacia. The Nkx6.1 homeobox sequence (including ~20 bp of flanking sequence at both sides) was PCR amplified with the primers 5′-CAAGAATTCATTTTGTTGGACAAAGATG-3′ (sense) and 5′-CTTCTCGAGGGTGGCCATCTCGGCT-3′ (antisense) containing an *Eco*RI and a *Xho*I site, respectively. The product was digested with *Eco*RI and *Xho*I and ligated into the expression vector pGEX-4T-1 at the *Eco*RI and *Xho*I sites so that the Nkx6.1 homeobox sequence was fused in frame to the 3′-end of the glutathione *S*-transferase (GST) gene. Both strands of the insert were sequenced to ensure that no PCR-generated mutations were incorporated. The GST-Nkx6.1HD fusion protein was synthesised in the protease-deficient *Escherichia coli* BL21 strain and purified on a glutathione Sepharose 4B column according to the manufacturer's manual.

2.6. Electromobility shift assays (EMSAs)

The reactions were carried out at 4°C in an EMSA buffer with a final concentration of 20% glycerol, 25 mM Tris-HCl pH 7.5, 150 mM KCl, 60 mM NaCl, 2 mM $MgCl_2$ and 1 mM DTT. Assays with nuclear extracts from cells were added 1 μg poly(dIdC) and 1 μg poly(dGdC), and assays with GST-Nkx6.1HD fusion protein were added 0.5 μg poly(dIdC), 0.5 μg poly(dGdC) and 1 μg bovine serum albumin (BSA). When antisera were used for supershifts, 1 μl undiluted antiserum was added. Between 0.02 and 2 pmol ^{32}P-labeled probe was added and incubated for 10 min at 4°C. Protein-DNA complexes were separated from unbound probe by 4 h electrophoresis on a 6% native polyacrylamide gel at 180 V and 4°C in 0.5×TBE buffer. The A45.A probe was generated from oligo A45.A: 5′-GAAGATCTCATTTAATTGGTTTTTAATTGGATCCGAGACTGAGCG-3′ (the two Nkx6.1 binding sites are underlined) and ^{32}P-phosphorylated antisense primer BSS3.A: 5′-CGCTCAGTCTCGGATCC-3′ and the complementary strand was filled out with Klenow. PCR probes were generated as described below.

2.7. Binding site selection

The Nkx6.1-binding site was determined in two independent experiments performed in parallel. A synthetic randomised (N_{20}) oligonucleotide probe was generated from oligo BSS1: 5′-CGGGCTGAGATCAGTCTAGATCTNNNNNNNNNNNNNNNNNNNNGGATCCGAGACTGAGCGTCGTC-3′ and ^{32}P-phosphorylated antisense primer BSS3: 5′-GACGACGCTCAGTCTCGGATCC-3′ and the complementary strand was filled out with Klenow. The binding reaction was carried out in 25 μl with 7 nM ^{32}P-labeled (N_{20}) oligonucleotide probe in EMSA buffer with 0.5 μg poly(dIdC), 0.5 μg poly(dGdC), 1 μg BSA, and 25 nM GST-Nkx6.1HD. The samples were incubated for 15 min at 4°C and separated on a 6% EMSA gel. Bound oligonucleotides were excised from the undried gel and one fifth was used directly in a PCR reaction to amplify the selected sequences. The PCR reaction was carried out in 50 μl (45 μl+the gel piece) with 20 pmol sense primer BSS2: 5′-CGGGCTGAGATCAGTCTAGATCT-3′, 20 pmol ^{32}P-phosphorylated antisense primer BSS3, 250 nM dNTPs, 2 units of DynaZyme II DNA polymerase from Finnzymes Oy and the buffer recommended by the manufacturer. The PCR conditions were 3 min at 95°C followed by 20 cycles of 50 s at 95°C, 50 s at 64°C, 1 min at 72°C. A binding reaction was carried out as described above with 5 μl of the PCR reaction as probe. The binding site selection and amplification procedure was repeated five times. The resulting selected sequences of the final PCR were phosphorylated and cloned in a blunt ended pBluescript SK II$^+$ vector from Stratagene. Forty-eight clones from each experiment were sequenced using a ^{33}P-labeled T7 primer and Thermo Sequenase from Amersham.

2.8. Transient transfection of NIH-3T3 cells

The Nkx6.1HD/VP16AD fusion construct was excised from RCAS-rNkx6.1-Homeo-VP16 by a *Cla*I digest and transferred to the *Cla*I site of pCMV4 [13]. The construction of RCAS-rNkx6.1-Homeo-VP16 will be described elsewhere. The reporter constructs have a pGL3-basic (from Promega) backbone with the 754 bp thymidine kinase (TK) promoter inserted in the *Bgl*II and the *Hin*dII sites upstream of the luciferase reporter gene. The oligo probe A45.A, which contain two Nkx6.1-binding sites and a *Bgl*II and a *Bam*HI site at the ends, was made double-stranded with the BSS3.A antisense primer and Klenow. The ends were cut with *Bgl*II and *Bam*HI and two copies were inserted at the conserved *Bgl*II site upstream of the TK promoter of the vector. As a control a similar reporter construct was generated with the oligo probe A45.B: 5′-GAAGATCTCATTTCCTTGGTTTTTCCTTGGATCCGAGACTGAGCG-3′, where the two Nkx6.1-binding sites are mutated (underlined) to abolish binding.

For the transient transfection assays we used 3×10^4 NIH-3T3 cells per well in 24 well plates from Nunc Inc. To each well of cells were added 2.5 ng pRL-CMV vector from Promega as internal standard, 400 ng reporter DNA, the indicated amount of effector DNA (pCMV4-Nkx6.1 (1.4 kb) (see below) or pCMV4-Nkx6.1HD/VP16AD), and filled up with empty pCMV4 vector to a total of 500 ng DNA per well. The DNA was incubated for 45 min at room temperature with 3 μl LipofectAMINE in 50 μl OptiMEM from Gibco-BRL, added to the cells and incubated. Then the cells were cultivated for 36 h, harvested and assayed for firefly and *Renilla* luciferase activity using the Dual-Luciferase Reporter Assay System from Promega.

2.9. Stable transfection of MSL-G-AN cells

For the stable expression of Nkx6.1 in the MSL-G-AN cells most of the long UTRs of the 3211 bp Nkx6.1 cDNA sequence were removed at the unique restriction sites *Fsp*I and *Apa*LI. The resulting Nkx6.1 cDNA piece of 1456 bp was blunt end ligated into the *Eco*RV site of the pBluescript SK II$^+$ vector from Stratagene and then inserted in the correct orientation at the *Hin*dIII and *Xba*I sites of the pCMV4 vector [13]. 2×10^8 MSL-G-AN cells were resuspended in 500 μl 1×PBS with 40 μg pCMV4-Nkx6.1 (1.4 kb) and 10 μg ploxPneo-1 vector (a kind gift from A. Nagy) conferring resistance to Geneticin (antibiotic G418 sulphate) from Gibco-BRL. The cells were electroporated as described [15] except that the cells were transferred to a 14 cm dish and cultivated for 24 h in normal medium before selection in medium containing 200 μg/ml G418. After 1 week the selection medium contained 100 μg/ml G418 until G418-resistant single cells had grown to form colonies. Eleven clones were isolated and cultivated separately. In vivo passage and reestablishment as cell lines was performed as described [16,17].

For RNA isolation 1×10^7 cells was extracted with 1 ml RNAzol B from Campro Scientific as described by the manufacturer. cDNA synthesis was carried out in 25 μl reactions as described [4]. Multiplex RT-PCR was performed as described in [4]. Twenty-five cycles of PCR were run for Nkx6.1 (284 bp) and G6PDH (214 bp), and 22 cycles were run for insulin (312 bp) and G6PDH (214 bp).

Nuclear extracts were prepared from 2×10^8 cells as described in [18] except that PMSF was replaced with AEBSF. Western blot analysis to test the Nkx6.1 protein levels in the clones was performed as described [14].

2.10. Cell lines

COS7 cells were cultivated in DMEM 4500 mg/l glucose, NIH-3T3 cells in DMEM 1000 mg/l glucose, both media containing 10% foetal bovine serum, 100 U/ml penicillin and 100 μg/ml streptomycin. MSL-G-AN cells were established and cultivated as described [17], NHI-6F-INS cells as described [16], and the AN+Pdx1 clone 1.1.10 as described [19]. All cell lines were kept in humidified incubators with 5% CO_2 at 37°C.

2.11. Immunocytochemical stainings and antibodies

Immunocytochemical stainings of monolayer cells were carried out with the Histostain-Plus kit from Zymed Laboratories Inc. Monoclonal antibodies against insulin (HUI-18) and glucagon (GLU-001) were from Novo-Nordisk Biolabs. A series of eight different antisera (1852–1859) were raised against a GST-Pdx1 fusion protein [19,20]. Rabbit antisera against Nkx6.1 were described in [4]. For immunocytochemistry α-Nkx6.1-174 and α-Pdx1-1856 were used. EMSA supershift antisera: α-Nkx6.1-173, α-Pdx1-1852. Western blot antiserum: α-Nkx6.1-173.

3. Results

3.1. Isolation and in vitro expression of the rat Nkx6.1 cDNA

Using primers derived from the hamster Nkx6.1 sequence a 220 bp fragment of the 3′-end of the coding sequence was PCR amplified from rat islet cDNA. This fragment was used as a probe in a high stringency colony hybridisation screening of a cDNA library made from growth hormone-stimulated newborn rat islets [12]. A 3211 bp cDNA containing a 1095 bp open reading frame that matched the hamster Nkx6.1 sequence was isolated. The insert contained 788 bp 5′-UTR and 1328 bp 3′-UTR. In the 5′-UTR 311 bp upstream of the predicted translation start codon another 'ATG' codon is found, but this is closely followed by a termination codon and would give rise to a polypeptide no longer than six amino acid residues (Fig. 1). The predicted 365 amino acid rat Nkx6.1 peptide sequence deviates by only 11 and six amino acid residues from the human and the hamster homologues, respectively (Fig. 2). The homeodomain sequences are identical between the three mammalian species and so are the NK decapeptides except for an asparagine in the human sequence where the rat and the hamster sequences have aspartate (Fig. 2).

Notably, [^{35}S]methionine labeling by in vitro transcription

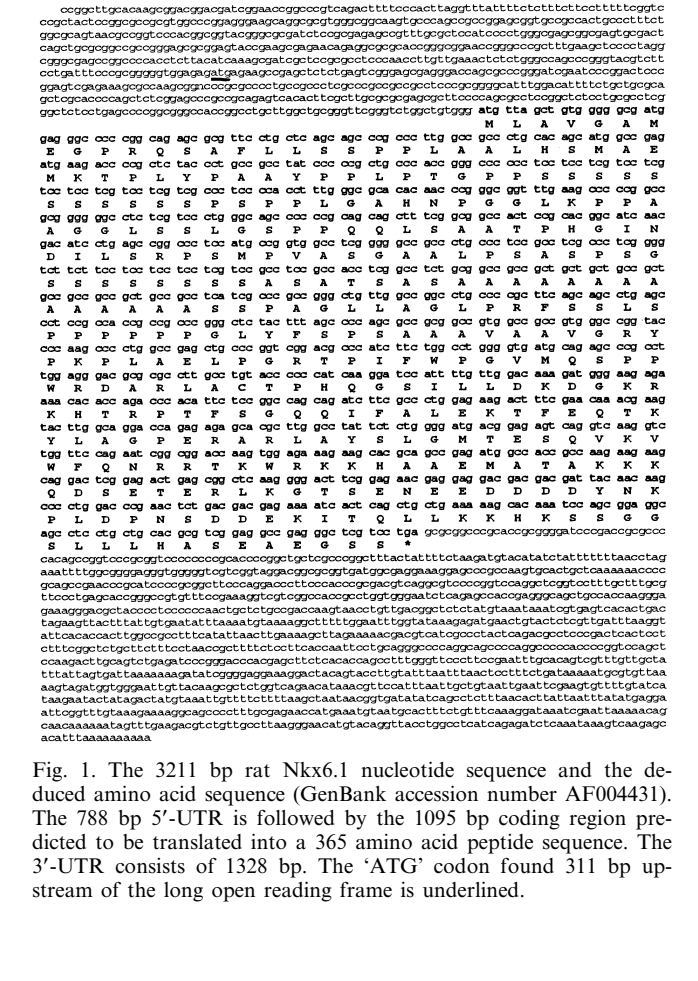

Fig. 1. The 3211 bp rat Nkx6.1 nucleotide sequence and the deduced amino acid sequence (GenBank accession number AF004431). The 788 bp 5′-UTR is followed by the 1095 bp coding region predicted to be translated into a 365 amino acid peptide sequence. The 3′-UTR consists of 1328 bp. The 'ATG' codon found 311 bp upstream of the long open reading frame is underlined.

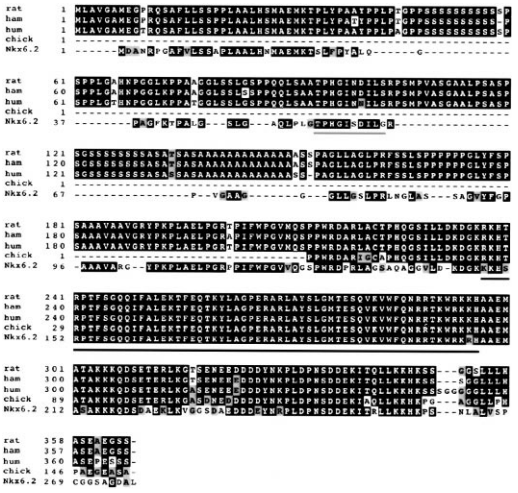

Fig. 2. A ClustalW alignment of the rat Nkx6.1 peptide sequence with the hamster, the human and the chicken (the sequence is from a partial clone) homologues. The mouse sequence of Nkx6.2/Gtx is also included. Identical positions are marked with black boxes and conservative substitutions are boxed in grey. The homeodomain is underlined with a black bar and the conserved NK decapeptide sequence is underlined with a grey bar. The Nkx6.2/Gtx sequence is closely related to Nkx6.1 and differs primarily in the lack of the long N-terminal serine and alanine stretches.

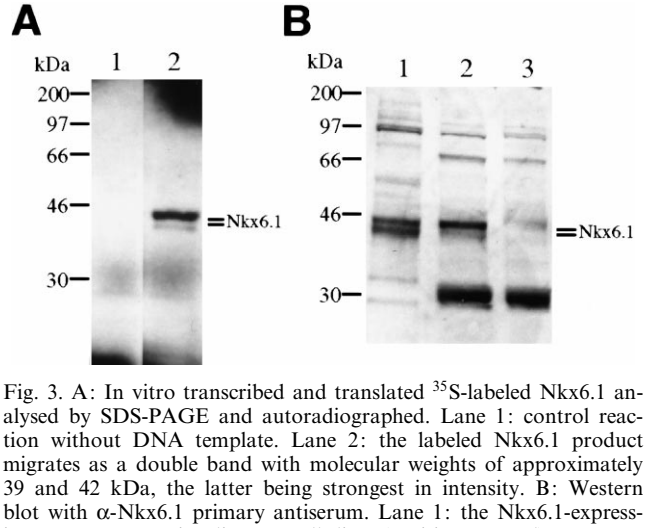

Fig. 3. A: In vitro transcribed and translated ^{35}S-labeled Nkx6.1 analysed by SDS-PAGE and autoradiographed. Lane 1: control reaction without DNA template. Lane 2: the labeled Nkx6.1 product migrates as a double band with molecular weights of approximately 39 and 42 kDa, the latter being strongest in intensity. B: Western blot with α-Nkx6.1 primary antiserum. Lane 1: the Nkx6.1-expressing NHI-6F-INS insulinoma cell line (positive control). Lane 2: Nkx6.1-transfected COS7 cells. Lane 3: COS7 cells (negative control).

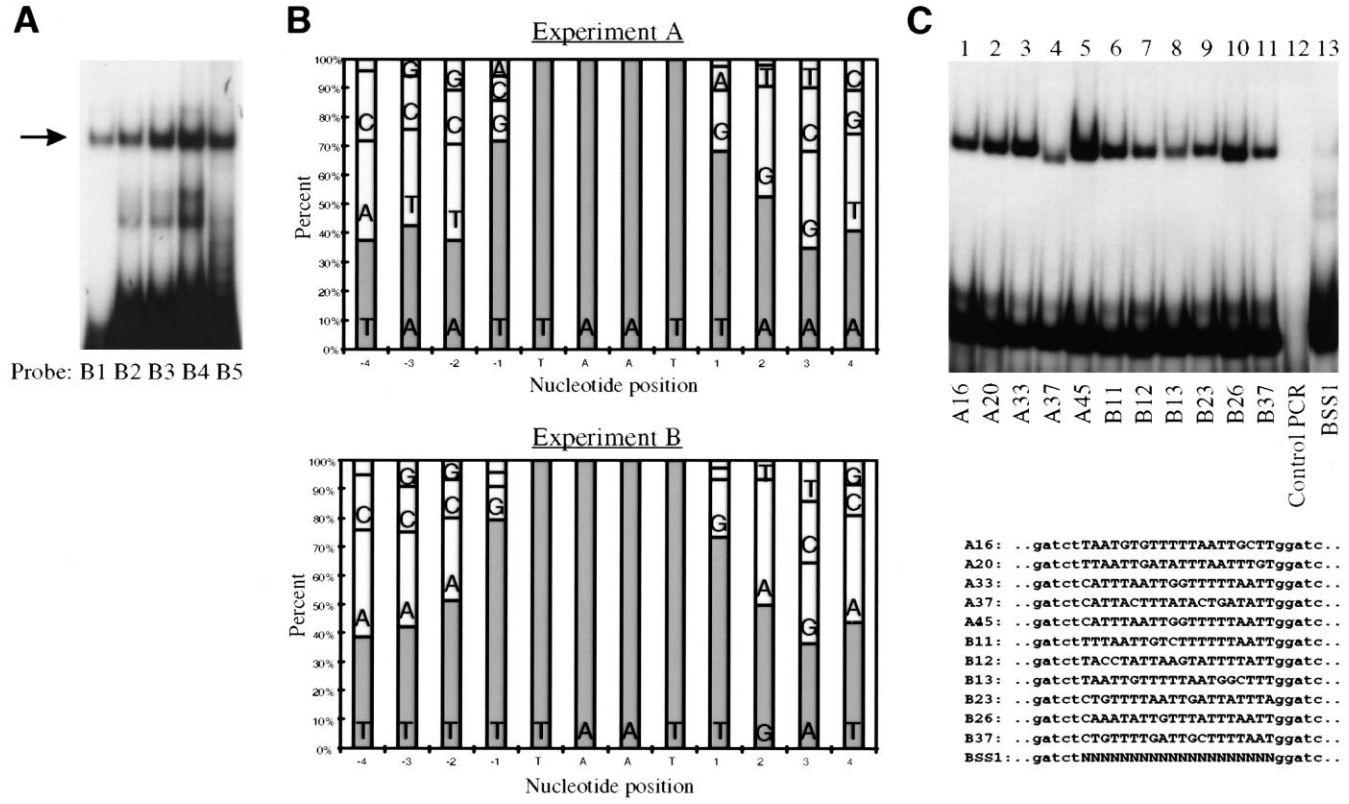

Fig. 4. A: The binding affinity of GST-Nkx6.1HD to the oligonucleotide pool (indicated by the arrow) increased for each cycle of binding site selection. Oligonucleotides that had been selected for binding to GST-Nkx6.1HD were PCR amplified and used as probes in an EMSA with 5 ng GST-Nkx6.1HD per lane. Lane B1: experiment B after one cycle of selection. Lane B2: experiment B after two cycles of selection and so forth. A similar result was obtained for experiment A. B: Alignment of all the TAAT motifs and the four flanking nucleotides at either side of the sequences selected for binding to GST-Nkx6.1HD. The results of the two independent experiments A and B were very similar. The DNA-binding site consensus sequence for Nkx6.1 is TTAATTG/A. C: Different binding site selected sequences were PCR amplified and show different binding affinities to 5 ng GST-Nkx6.1HD in an EMSA. Lanes 1–5: sequences from experiment A. Lanes 6–11: sequences from experiment B. Lane 12: control PCR without template. Lane 13: the randomised (N_{20}) oligonucleotide, BSS1.

and translation of the isolated cDNA resulted in two products with approximate molecular weights of 42 and 39 kDa (Fig. 3A). Using an α-Nkx6.1 antiserum [4] we performed Western blot analysis of COS7 cells transfected with the isolated Nkx6.1 cDNA. As seen in Fig. 3B two bands of molecular weights 42 and 39 kDa were detected, confirming the result obtained by [^{35}S]methionine labeling.

3.2. Nkx6.1 binds a specific DNA consensus sequence

In order to identify the optimal DNA-binding sequence for Nkx6.1, a glutathione *S*-transferase Nkx6.1 homeodomain (GST-Nkx6.1HD) fusion protein was synthesised and used in a PCR-based binding site selection on randomised synthetic oligonucleotides. Two independent experiments (called A and B) of five cycles of selection were performed in parallel and as shown in Fig. 4A the binding affinities increased markedly for each cycle. Finally, the resulting oligonucleotides were cloned and 48 colonies from both selection A and B were randomly picked and the inserts sequenced. Every single insert contained at least one TAAT motif. In experiment A a total of 120 TAAT sequences were identified giving an average of approximately three per insert, and the result of the alignment including the four flanking nucleotides at either side is shown in Fig. 4B. Experiment B gave a very similar result (Fig. 4B), and the deduced DNA-binding site consensus sequence for Nkx6.1 obtained in vitro is TTAATTG/A.

A number of the selected sequences were tested for binding to GST-Nkx6.1HD and as exemplified in Fig. 4C they show different binding affinities depending on how closely the sequences match the consensus and also depending on the number of binding sites they contain. The strongest binding sequence is A45, which contains two perfect consensus binding sites, whereas A37 with only one TAAT motif and no flanking nucleotides that match the consensus is a poor binder. Other strong binding sequences are A33 (which is identical to A45), A20 and B11 which also contain two closely matching binding sites, and B26 with one perfectly matching Nkx6.1-binding site.

To test the authenticity of the selected binding site consensus sequence, nuclear extracts from glucagonoma cells stably expressing Nkx6.1 (see below) were used in an EMSA with the A45 sequence as probe. This confirmed that TTAATTG is a binding site for the exogenously expressed full-length Nkx6.1 protein, which forms complexes with the A45 probe recognised by the α-Nkx6.1 antiserum (Fig. 5). Another β-cell homeodomain factor, Pdx1, also binds the A45 probe (Fig. 5). Binding of Pdx1 is not surprising given the close resemblance of the A45 binding sequence TTAATTG to the canonical Pdx1 binding site C/TTAATG [21,22]. EMSA analysis with mutated binding sites showed that the nucleotides flanking the TAAT core sequence are important for optimal binding of full-length Nkx6.1 (data not shown).

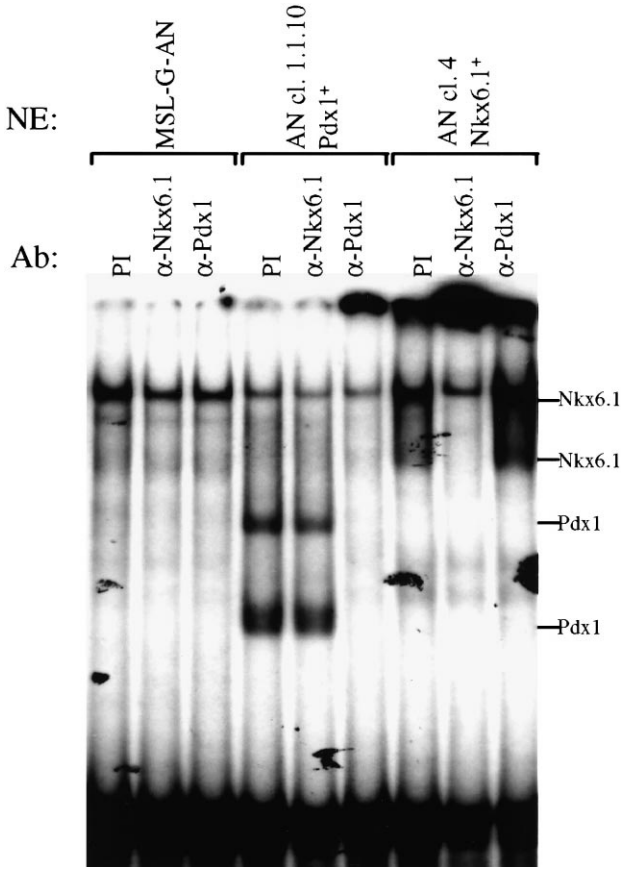

Fig. 5. Full-length Nkx6.1 binds the TTAATTG sequence. However, another homeodomain factor, Pdx1, can also bind this probe. EMSA with 5 μg nuclear extracts (NE) from MSL-G-AN cells, a stable Pdx1-expressing MSL-G-AN clone, and a stable Nkx6.1-expressing MSL-G-AN clone. Supershifts with the antisera (Ab) α-Pdx1 and α-Nkx6.1. PI: preimmune serum. Probe: double-stranded ^{32}P-labeled A45.A (containing two TTAATTG sites).
←

3.3. *Transactivating effect of Nkx6.1HD/VP16AD fusion construct*

To investigate if the binding of Nkx6.1 to the isolated consensus sequence could take place in an in vivo-like situation, we set up a transfection experiment in NIH-3T3 cells. Here the transactivation ability of a chimeric protein with the Nkx6.1 homeodomain sequence fused to the VP16 activation domain (Nkx6.1HD/VP16AD) was tested on a reporter plasmid containing four copies of the selected Nkx6.1-binding site (TTAATTG) inserted in front of the TK promoter. As a control an equivalent TK promoter with four mutated Nkx6.1-binding sites (TTCCTTG) inserted at the same position was used.

The TK promoter sequence itself was subsequently identified to contain two TAAT motifs (GTAATAT and TTAATAT) which bind the Nkx6.1HD with some but not optimal affinity. Consequently, the Nkx6.1HD/VP16AD activates the TK promoter to some extent (7–10-fold), and the induction level of the TK promoter alone with no inserts is similar to the induction level for the TK promoter with the mutated binding sites inserted (data not shown). However, when four copies of the Nkx6.1 binding site are inserted in front of the

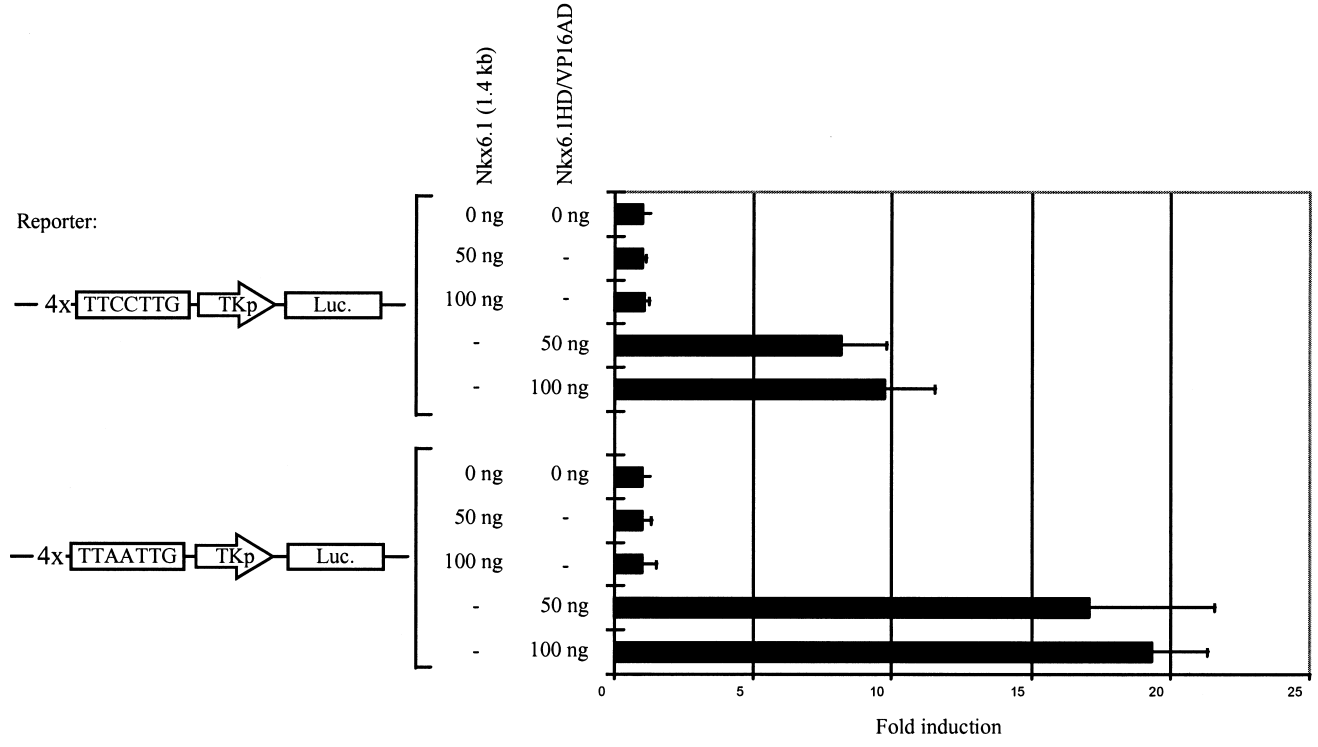

Fig. 6. The TTAATTG sequence functions as a binding site for Nkx6.1HD in a transient transfection experiment in NIH-3T3 cells. Four TTAATTG Nkx6.1-binding sites were inserted in front of the TK promoter driving the luciferase reporter gene. The control had four mutated Nkx6.1-binding sites (TTCCTTG) inserted at the same position. Full-length Nkx6.1 is not capable of affecting the TK promoter activity, whereas the Nkx6.1HD/VP16AD fusion construct activates both reporters due to two TAAT motifs within the TK promoter sequence. The induction of the promoter with the four Nkx6.1-binding sites is twice as strong as the induction of the control promoter with the mutated sites. The average of three independent experiments is presented, error bars indicate S.E.M.

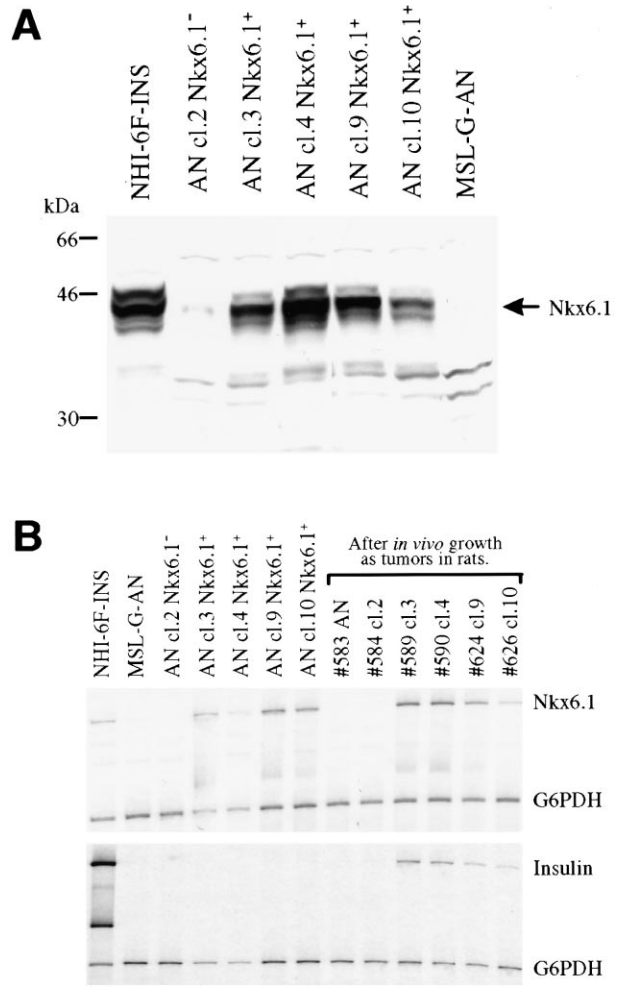

Fig. 7. A: Western blot analysis of the Nkx6.1 protein levels in the MSL-G-AN clones. B: Multiplex RT-PCR analysis showing induction of insulin expression in Nkx6.1-positive clones after in vivo passage.

TK promoter, the activation is approximately 20-fold (Fig. 6). This shows that the TTAATTG sequence is recognised by the Nkx6.1HD in an in vivo context. In a similar experiment full-length Nkx6.1 is not capable of activating the TK promoter (Fig. 6).

3.4. Screening for potential Nkx6.1 target genes

To further investigate the function of Nkx6.1, the Nkx6.1 cDNA under transcriptional control of the CMV promoter was used to stably transfect an in vitro culture of MSL-G-AN tumour cells, an Nkx6.1-negative rat α-cell line that forms glucagonomas when passaged subcutaneously in rats [17,23]. Eleven G418-resistant MSL-G-AN clones were isolated. Four clones (clones 3, 4, 9 and 10) expressed high levels of Nkx6.1 protein (Fig. 7A) and these were chosen for further analysis. As control cells we used untransfected MSL-G-AN and the Nkx6.1-negative clone 2 (Fig. 7A,B).

Empirically we know that the MSL-G-AN cells can lose their hormone expression solely because of prolonged in vitro cultivation, so the clones were grown in vivo as tumours in rats [17], and were subsequently screened by multiplex RT-PCR for changes in transcription of a number of islet cell-specific genes. We did not observe any changes in the expression level of the hormones glucagon, somatostatin and IAPP, but after the in vivo passage a small percentage (~1%) of the cells had become insulin-positive in all the Nkx6.1-positive clones 3, 4, 9 and 10 but not in the Nkx6.1-negative clone 2 nor in MSL-G-AN cells (Figs. 7B and 8). The expression of the β-cell factors Pax4, Pdx1, GLUT2 and GLP-1-R was not induced by the presence of Nkx6.1 in these clones, and the expression level of other important islet cell factors such as CCK, Pax6, Nkx2.2, Isl1, NeuroD, HNF1α, HNF4α, glucokinase and Brn4 was not affected (data not shown). Double immunofluorescent stainings confirmed the absence of Pdx1 immunoreactivity in the insulin-producing sub-population (data not shown).

4. Discussion

Here we have cloned the rat Nkx6.1 cDNA sequence, identified the DNA-binding site to be TTAATTG/A and shown that it is recognised in vivo by the Nkx6.1 homeodomain when present in a reporter plasmid. This is in agreement with the recent finding that Nkx6.2, which has a homeodomain sequence almost identical to Nkx6.1, binds with high affinity to a TAATTA sequence [24]. Nkx6.2 was initially shown to bind a MEF-2 motif (CTAAAAATAAC), which is the binding site for serum response factor-related proteins (RSRFs). Therefore Nkx6.2 was hypothesised to be a transcriptional repressor of serum-induced genes by competing with the RSRFs for binding to the same sites [6]. Despite the high degree of sequence identity between the homeodomains of Nkx6.1 and Nkx6.2, we have not been able to detect Nkx6.1 binding to the MEF-2 motif (M.C. Jørgensen, unpublished observation).

For most homeodomain factors a TAAT motif is the core sequence essential for binding to DNA (reviewed in [25]). One of the few exceptions is Nkx2.1, which preferentially binds a CAAG core [26]. By systematically combining different amino acid residue substitutions in the helix III and the N-terminal arm of the Nkx2.1 homeodomain with point mutations of the DNA-binding site, it was shown that particularly the amino acid residues 6, 7, 8, 50, 51 and 54 of the homeodomain sequence interact closely with the specific bases of the core nucleotide sequence [27]. It was thus shown that amino acids Q6, T7 and Y8 in the N-terminal end of the Antennapedia homeodomain would dictate the 5′-T in the core motif (TAAT), whereas V6, L7 and F8 found in the Nkx2.1 homeodomain dictate a C at that site (CAAG). The 3′-A (TAAT) is specified by amino acid residue 51 which is the most conserved among all homeodomains (almost always N), and M54 in the Antennapedia homeodomain would dictate the 3′-T in the core sequence (TAAT), whereas the Nkx2.1 homeodomain Y54 prefers a G at that position (CAAG). Finally, it was shown that the two nucleotides immediately 3′ of the TAAT core (TAATNN) are specified by amino acid residue 50. Thus, K50 in the bicoid homeodomain dictates CC at those sites, whereas Q50 in the Nkx2.1 homeodomain dictates TG. Based on our selected binding site, the N-terminal amino acids P6, T7 and F8 in the Nkx6.1 homeodomain sequence appear to prefer a T in the most 5′ position of the core motif (TAAT). Similarly, T54 in the Nkx6.1HD seems to prefer a T at the most 3′ position of the core (TAAT). The N51 and Q50 amino acids in the Nkx6.1 homeodomain are consistent with the nucleotides A found

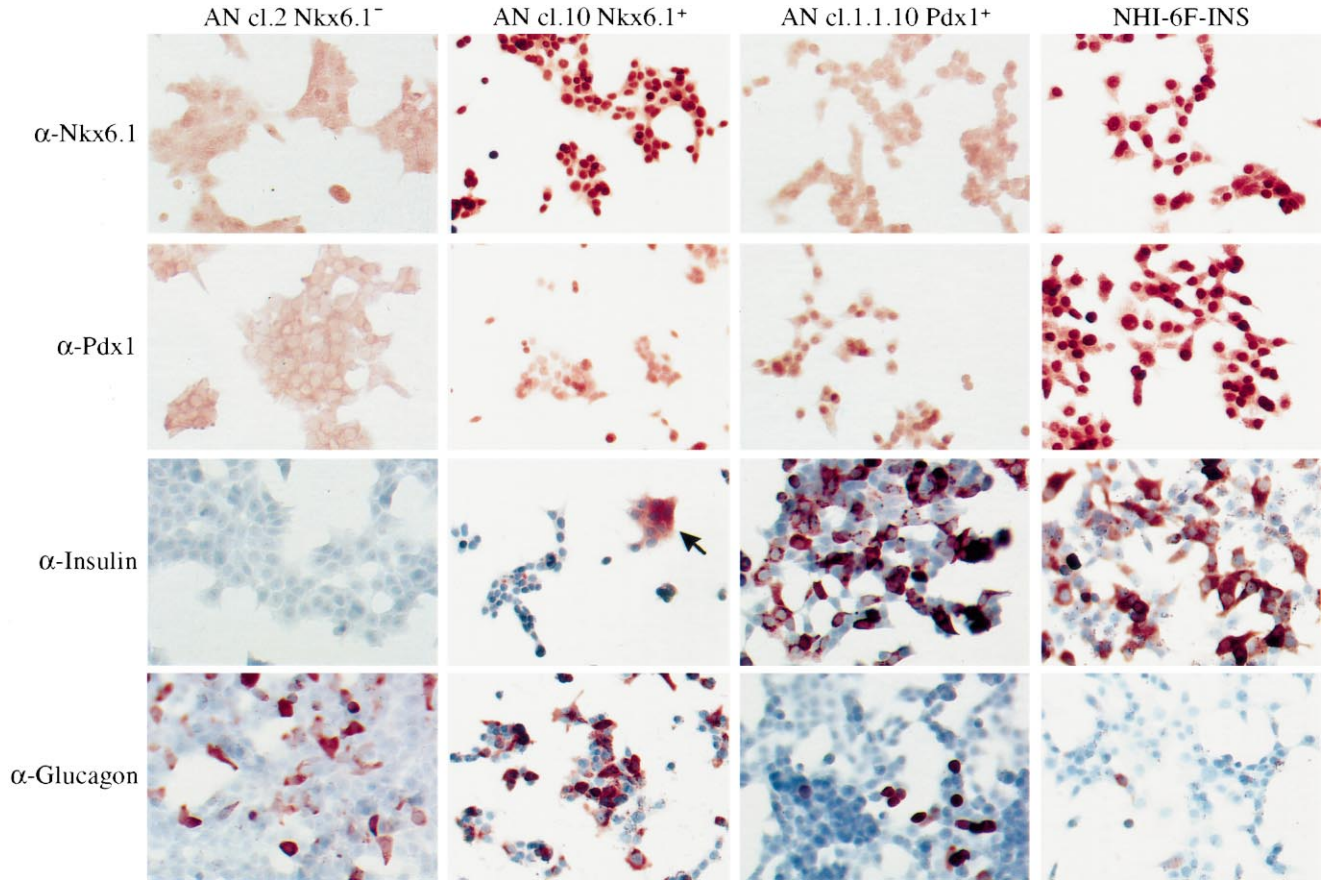

Fig. 8. Insulin production in Nkx6.1-expressing MSL-G-AN cells. Immunoperoxidase stainings of Nkx6.1$^-$ (clone 2), Nkx6.1$^+$ (clone 10), Pdx1$^+$ (clone 1.1.10) AN cells, and NHI-6F-INS cells with the indicated antisera. The Nkx6.1-expressing AN clones show induction of insulin expression in a subset of the cells (indicated by the arrow), whereas the glucagon expression levels are unaffected. The α-Nkx6.1 and α-Pdx1 antisera give distinct red nuclear stainings, whereas the α-insulin and α-glucagon antibody stainings (also in red) are cytoplasmic.

at position 3 and TG found at the 3′ flanking positions of the core in the selected sequence (TTAATTG/A).

In general, homeodomain proteins are known to interact with DNA with relatively low binding specificity compared to other classes of transcription factors (reviewed in [28]). As a consequence they do not show much binding specificity in vitro and bind indiscriminately to many TAAT-containing sequences. Increased specificity is often accomplished in vivo by co-operative binding with one or more cofactors [29]. An example of this is the complexes formed between Pbx and Prep proteins and members of the Hox family of HD proteins [30]. Notably, the sequence selected by the Nkx6.1HD is also recognised by Pdx1, another homeodomain protein important for pancreas development and β-cell function [10,31,32]. Absence of Nkx6.1 or Pdx1, which are co-expressed during most of the pancreatic development [2], have very different consequences for pancreas development (M. Sander, personal communication, [31,32]), suggesting that these proteins regulate distinct target genes. This question could be addressed by the use of different cofactors to achieve different DNA-binding specificities between the two proteins. Indeed, Pdx1 forms distinct cell type-specific complexes with Meis and Pbx proteins [33,34]. Such a scenario might make it difficult to identify the direct target genes for a homeodomain factor like Nkx6.1 based on in vitro selected binding sites. A new family of homeodomain-interacting protein kinases (HIPKs) has recently been identified, and HIPK2 was shown to enhance the DNA-binding activity of *Drosophila* NK-3 and thereby act as a co-repressor of NK-3 autorepression of its own promoter [35]. Potentially the HIPKs could be possible cofactor candidates for Nkx6.1 if they are expressed in the β-cells.

It has been suggested that Nkx6.1 could be a repressor of glucagon expression [10] and the glucagon promoter sequence contains a perfect Nkx6.1 consensus binding site at position −492 to −486 relative to the translational start site. However, in the stable Nkx6.1-expressing MSL-G-AN clones, we found the glucagon expression level unchanged after in vivo passage. Consistent with this, Nkx6.1-expressing αTC1.9 cells [1] show a high glucagon expression level. In transient transfection studies we have tested the effect of Nkx6.1 on the glucagon and insulin promoters, but failed to detect any effect on the transcriptional activities of these promoters in both islet and non-islet cell lines (M.C. Jørgensen, unpublished observations). Nevertheless, expression of Nkx6.1 in the MSL-G-AN cells combined with in vivo culture causes the induction of insulin in a sub-population of the cells, suggesting that Nkx6.1 can induce insulin production possibly by an indirect mechanism. The limited number of cells that produce insulin could reflect heterogeneity of the MSL-G-AN cells, where the presence of Nkx6.1 is sufficient in only a sub-population of the cells to trigger the shift towards a β-cell-like phenotype. In contrast to Nkx6.1, Pdx1 is sufficient to induce insulin production in MSL-G-AN cells without the need for an in vivo culture step [19]. However, we were not able to detect Pdx1

expression in Nkx6.1-transfected AN cells after in vivo culture suggesting either that transient expression of Pdx1 occurs or that a Pdx1-independent pathway of insulin gene activation is operating in these cells.

Acknowledgements: This work was supported by the Juvenile Diabetes Foundation International. M.C. Jørgensen was supported by the Danish National Research Fund. J. Ericson was supported by the Swedish Foundation for Strategic Research and The Swedish Natural Science Research Council. We thank Ragna Jorgensen and Tove F. Nielsen for expert technical assistance and Heidi I. Jensen and Erna E. Petersen for advice and instruction with regard to animal work and immunocytochemistry.

References

[1] Rudnick, A., Ling, T.Y., Odagiri, H., Rutter, W.J. and German, M.S. (1994) Proc. Natl. Acad. Sci. USA 91, 12203–12207.
[2] Øster, A., Jensen, J., Serup, P., Galante, P., Madsen, O.D. and Larsson, L.-I. (1998) J. Histochem. Cytochem. 46, 707–715.
[3] Øster, A., Jensen, J., Edlund, H. and Larsson, L.I. (1998) J. Histochem. Cytochem. 46, 717–721.
[4] Jensen, J., Serup, P., Karlsen, C., Nielsen, T. and Madsen, O. (1996) J. Biol. Chem. 271, 18749–18758.
[5] Qiu, M., Shimamura, K., Sussel, L., Chen, S. and Rubenstein, J.L. (1998) Mech. Dev. 72, 77–88.
[6] Komuro, I., Schalling, M., Jahn, L., Bodmer, R., Jenkins, N.A., Copeland, N.G. and Izumo, S. (1993) EMBO J. 12, 1387–1401.
[7] Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C.H., Ward, J.M. and Gonzalez, F.J. (1996) Genes Dev. 10, 60–69.
[8] Briscoe, J., Sussel, L., Serup, P., Hartigan-O'Connor, D., Jessell, T.M., Rubenstein, J.L. and Ericson, J. (1999) Nature 398, 622–627.
[9] Sussel, L., Kalamaras, J., Hartigan-O'Connor, D.J., Meneses, J.J., Pedersen, R.A., Rubenstein, J.L. and German, M.S. (1998) Development 125, 2213–2221.
[10] Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K. and Edlund, H. (1998) Genes Dev. 12, 1763–1768.
[11] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1998) Current Protocols in Molecular Biology, Vol. 1, pp. 6.1.1–6.3.6.
[12] Michelsen, B., Petersen, J., Boel, E., Møldrup, A., Dyrberg, T. and Madsen, O. (1991) Proc. Natl. Acad. Sci. USA 88, 8754–8758.
[13] Andersson, S., Davis, D.L., Dahlbäck, H., Jörnvall, H. and Russell, D.W. (1989) J. Biol. Chem. 264, 8222–8229.
[14] Møller, C., Christgau, S., Williamson, M., Madsen, O., Niu, Z., Bock, E. and Bækkeskov, S. (1992) Mol. Endocrinol. 6, 1332–1342.
[15] Serup, P. and Schøller, J. (1989) Diabetologia 32, 409–415.
[16] Madsen, O.D., Andersen, L.C., Michelsen, B., Owerbach, D., Larsson, L.-I., Lernmark, Å. and Steiner, D.F. (1988) Proc. Natl. Acad. Sci. USA 85, 6652–6656.
[17] Madsen, O. et al. (1993) Endocrinology 133, 2022–2030.
[18] Schreiber, E., Mathias, P., Muller, M. and Schaffner, W. (1988) EMBO J. 7, 4221–4229.
[19] Serup, P., Jensen, J., Andersen, F., Jorgensen, M., Blume, N., Holst, J. and Madsen, O. (1996) Proc. Natl. Acad. Sci. USA 93, 9015–9020.
[20] Peers, B., Leonard, J., Sharma, S., Teitelman, G. and Montminy, M. (1994) Mol. Endocrinol. 8, 1798–1806.
[21] Ohlsson, H., Karlsson, K. and Edlund, T. (1993) EMBO J. 12, 4251–4259.
[22] Petersen, H., Serup, P., Leonard, J., Michelsen, B. and Madsen, O. (1994) Proc. Natl. Acad. Sci. USA 91, 10465–10469.
[23] Madsen, O.D. et al. (1997) Horm. Metab. Res. 29, 265–270.
[24] Awatramani, R., Scherer, S., Grinspan, J., Collarini, E., Skoff, R., O'Hagan, D., Garbern, J. and Kamholz, J. (1997) J. Neurosci. 17, 6657–6668.
[25] Gehring, W.J., Affolter, M. and Bürglin, T. (1994) Annu. Rev. Biochem. 63, 487–526.
[26] Damante, G. et al. (1994) Nucleic Acids Res. 22, 3075–3083.
[27] Damante, G. et al. (1996) EMBO J. 15, 4992–5000.
[28] Mannervik, M. (1999) BioEssays 21, 267–270.
[29] Mann, R.S. and Chan, S.K. (1996) Trends Genet. 12, 258–262.
[30] Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F. and Blasi, F. (1998) EMBO J. 17, 1434–1445.
[31] Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L. and Wright, C.V. (1996) Development 122, 983–995.
[32] Jonsson, J., Carlsson, L., Edlund, T. and Edlund, H. (1994) Nature 371, 606–609.
[33] Peers, B., Sharma, S., Johnson, T., Kamps, M. and Monteminy, M. (1995) Mol. Cell. Biol. 15, 7091–7097.
[34] Swift, G.H., Liu, Y., Rose, S.D., Bischof, L.J., Steelman, S., Buchberg, A.M., Wright, C.V. and MacDonald, R.J. (1998) Mol. Cell. Biol. 18, 5109–5120.
[35] Kim, Y.H., Choi, C.Y., Lee, S.J., Conti, M.A. and Kim, Y. (1998) J. Biol. Chem. 273, 25875–25879.