The caudal-related Homeodomain Protein Cdx-2/3 Regulates Glucagon Gene Expression in Islet Cells*

(Received for publication, January 26, 1996, and in revised form, July 1, 1996)

Beate Laser‡§, Paolo Meda¶, Isabel Constant‡, and Jacques Philippe‡

From ‡Clinical Diabetology, Department of Medicine and the ¶Department of Morphology, Centre Médical Universitaire, 1 rue Michel Servet, CH-1211 Geneva 4, Switzerland

Glucagon gene transcription in the endocrine pancreas is regulated by at least four cis-acting DNA control elements. We showed previously that G1 is critical for alpha cell-specific expression. G1 contains three AT-rich sequences important for promoter function, which represent candidate binding sites for homeodomain transcription factors. Performing reverse transcription-polymerase chain reaction amplifications with degenerate oligonucleotide primers homologous to the Antennapedia homeobox, cDNA clones corresponding to the caudal-related gene cdx-2/3 were predominantly obtained from glucagon-producing cells and primary non-beta cells. From RNase protection and polymerase chain reaction analyses, cdx-2/3 turned out to be the only caudal-related gene that is expressed at significant levels in cells of the endocrine pancreas. Cdx-2/3 binds with high affinity to an AT-rich motif of G1, which matches the consensus binding site of caudal-related proteins. In the glucagon-producing hamster cell line InR1G9, Cdx-2/3 is a subunit of complex B3 formed on G1. Alternative splicing generates two cdx-2/3 transcripts in islet cells, coding for a full-length protein and an amino-terminally truncated isoform. Although both isoforms bind G1 with similar affinity, only the full-length Cdx-2/3 A protein activates glucagon gene transcription in non-glucagon-producing cells, transcriptional activation being dosedependent. We therefore conclude that the caudal-related gene cdx-2/3 is implicated in the transcriptional control of glucagon gene expression in the alpha cells of the islets of Langerhans.

Glucagon and the glucagon-like peptides are synthesized as a common precursor, preproglucagon, encoded by the glucagon gene and are involved in the control of glucose homeostasis. Expression of the glucagon gene is highly restricted to the alpha cells of the endocrine pancreas, the L cells of the intestine, and certain areas of the brain (1–3). The factors controlling glucagon gene expression are poorly understood. In addition, they may differ depending on the tissue examined; studies in transgenic mice expressing the simian virus (SV) large T antigen under the control of 1300 bp of the rat glucagon gene 5′-flanking sequences demonstrate expression in the pancreas and the brain, whereas additional upstream sequences are necessary for expression in the gut (4, 5).

Tissue-specific expression of the glucagon gene in the pancreas is conferred in two steps by the islet-specific enhancer elements G2, G3, and G4 (6–8) and the alpha cell-specific proximal promoter element G1 (7, 9, 10). G1 contains three AT-rich sequences that are candidate binding sites for homeodomain transcription factors. Two of these sequences represent nearly identical 7-bp direct repeats and are important for transcriptional activity. At least four protein complexes were found to interact with G1, and the integrity of the AT-rich 7-bp direct repeats was shown to be critical for their binding (9).

In a first attempt to identify the proteins implicated in the control of glucagon gene expression, we designed degenerate oligonucleotide primers based on highly conserved sequences in the first and third helix of the Antennapedia class of homeodomain transcription factors. Recent studies indicate that homeobox genes are involved in the control of pancreatic hormone genes, such as insulin and somatostatin. Several homeodomain proteins have been isolated from islet cells by cDNA cloning or PCR amplification of reverse transcribed RNA (11, 12) and proposed to be implicated in insulin and somatostatin gene regulation (13–16).

In this study, we show that the caudal-related gene cdx-3, which has been proposed to regulate insulin gene transcription through its binding to the FLAT element (15) is expressed in glucagon-producing cells. Since cdx-3 represents the hamster homologue of the mouse cdx-2 gene (15, 17, 18), this gene will be referred to here as cdx-2/3. In cell lines of the endocrine pancreas, alternative splicing generates two cdx-2/3 mRNAs, encoding a full-length protein (termed Cdx-2/3 A) and an amino-terminally truncated isoform (Cdx-2/3 B). Cdx-2/3 binds the proximal AT-rich direct repeat motif of the glucagon promoter element G1, which matches the consensus binding site of caudal-related genes (18, 19) and is contained in complex B3 binding to G1. Interestingly, the full-length Cdx-2/3 A, but not the truncated Cdx-2/3 B isoform activates glucagon gene expression in non-glucagon-producing cells. Different transactivating properties of the two isoforms exhibiting the same affinity for G1 suggests that differential expression of cdx-2/3 A and cdx-2/3 B may regulate transcriptional activation. We conclude that the intestine-specific gene cdx-2/3 is the only caudal-related gene expressed in the endocrine pancreas and that it may be involved in the control of pancreatic glucagon gene expression.

*This work was supported by Swiss National Science Foundation Grant FN 3200-031227.91 (to J. P.) and Grant 32-043086.95 (to P. M.), the Helmut Horten Foundation, the Institute for Human Genetics and Biochemistry, the Wolfermann-Nageli Foundation, and the Elsie and the Helmut Horten Foundation, the Institute for Human Genetics and Biochemistry, the Wolfermann-Nageli Foundation, and the Elsie and Carlos de Reuter Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom all correspondence should be addressed: Clinical Diabetology, Centre Medical Universitaire, 1 rue Michel Servet, CH-1211 Geneva 4, Switzerland. Tel.: +41 22 702 55 67; Fax: +41 22 702 55 43; E-mail: Laser@cmu.unige.ch.

¶ The abbreviations used are: SV, simian virus; bp, base pair(s); PCR, polymerase chain reaction; nt, nucleotide(s); BHK, baby hamster kidney; EMSA, electromophoretic mobility shift assay; CAT, chloramphenicol acetyltransferase; RSV, Rous sarcoma virus.
Glucagon Gene Expression

TABLE I

| Glucagon-producing cell lines | Insulin-producing cell lines | Primary islets | Primary non-beta cells | Primary beta cells | Reference |
|------------------------------|-----------------------------|---------------|------------------------|-------------------|----------|
| Hox A1                       | 1/23                        | 1/28          | 12/12                  | 7/11              | 44       |
| Hox A9                       | 2/23                        | 1/28          | 24/28                  | 7/11              | 45       |
| Hox C10                      | 1/23                        | 2/28          | 22/28                  | 10/10             | 46       |
| Cdx-2/3                      | 16/23                       | 12/12         | 11/16                  | 43                |         |
| IDX1/STF1/IPF1               | 7/11                        | 1/28          | 15/17                  | 49*               |         |
| Unknown                      |                             |               |                        |                   |          |

EXPERIMENTAL PROCEDURES

Isolation of Islets of Langerhans and Separation of Beta and Non-Beta Cells—Rat pancreatic islets as well as primary beta and non-beta cells were isolated as described previously (20, 21).

Reverse Transcription and PCR—cDNA was generated from 2 μg of total RNA from different islet cell lines of various origins. Amplification reactions were carried out for 40 cycles according to standard protocols (22). Degenerate primers used for PCR amplification of homeodomain proteins were as follows: 5'-GA(G/A)CT/TT/C/G/GA(G/A)AA/AA/AG/G A/TT-3' and 5'-CAT/C/G(IG/C/G)/AT/G/T/C/T/G/TT/G/AA/ACCA-3' (nt 43–59 and 142–162 of the Antennapedia class homeobox, respectively). PCR fragments were sequenced by the double-stranded dideoxy chain termination method (22), and sequence comparisons were performed at the National Center for Biotechnology Information (NCBI) using the BLAST Network Service.

Cell Culture and DNA Transfection—The glucagon-producing InR1G9 (hamster; Ref. 24) and αTC1 (mouse; Ref. 25), insulin-producing HIT T15 (hamster; Ref. 26), βTC1 (mouse; Ref. 27), RIN5-F (rat; Ref. 28), and MIN6 (mouse; Ref. 29), and somatostatin-producing RIN 1027-B (rat; Ref. 30) cell lines as well as the non-β islet Syrian hamster baby kidney (BHK 21) cells were grown in RPMI 1640 (Seromed; Basel, Switzerland) supplemented with 5% fetal calf serum and 5% newborn calf serum, 100 units of penicillin, and 100 μg of streptomycin/ml. The insulin-producing rat cell line INS-1 (31) was cultured in RPMI 1640 supplemented with 10% fetal bovine serum, 5% heat-inactivated fetal calf serum, 10 units of penicillin, and 100 μg of streptomycin/ml. αTC1 and βTC1 cells were generously provided by Dr. D. Hanahan (University of California, San Francisco), MIN6 cells by Dr. J.-I. Miyazaki (Kumamoto University, Japan), and INS1 cells by Dr. C. B. Wolfe (University of Geneva Medical School, Geneva, Switzerland). BHK cells were transfected by the calcium phosphate technique (32) using 10 μg of plasmid DNA/6-cm Petri dish. The cdx-3 expression plasmid pBAT7.cdx3 was kindly provided by Dr. M. German (University of California, San Francisco, CA), and plasmids cdx2/3 A.cmv, cdx2/3 B.cmv, and cdx4.cmv were constructed by subcloning the respective PCR fragments in a cmv-driven expression vector (33). For stable transfection of InR1G9 cells, the plasmid fragment of cdx-2/3 B was cloned in sense orientation into the eukaryotic expression vector pBJ1-neo (Ref. 34; kindly provided by Dr. M. Davies, Stanford University School of Medicine, Stanford, CA). After transfection by the calcium phosphate technique, stably transfected clones were selected and grown in 200–400 μg/ml Geneticin (G-418 Sulfate, Life Technologies, Inc., Basel, Switzerland).

RNAase Protection Analysis—Uniformly labeled RNA probes were obtained from in vitro transcription with T3 or T7 RNA polymerase (Boehringer Mannheim) as described (35). The cdx-1 cDNA used for riboprobe generation was generously provided by Dr. R. James (University of Melbourne, Melbourne, Australia) and contains 412 nt homologous to nt 114–525 of the mouse cdx-1 mRNA (36). cDNAs for the mouse cdx-4 and the hamster cdx-2/3 A, cdx-2/3 B, and cdx-2/3 homeobox were obtained by PCR amplification. As a control for RNA quantity, a riboprobe was synthesized containing 250 nt of the mouse β-actin cDNA (codon 220–303; Ambion,lagano, Switzerland). RNAse protection analysis was performed according to standard protocols (22) except that RNA digestion was performed with 4 μg/ml RNase A. The relative abundance of protected fragments was analyzed using a PhosphorImager.

Electrophoretic Mobility Shift Assays (EMSAs)—EMSAs were performed as described previously (37) using nuclear extracts prepared according to Schreiber et al. (38). Anti-Cdx-3 antibodies were generously provided by Dr. M. German (University of California).

Immunoblotting—Nuclear extracts were separated on 15% SDS-polyacrylamide gels and transferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA) as described previously (39, 40). Western blotting was performed using rabbit polyclonal anti-Cdx-3 (15) as first antibody and peroxidase-coupled anti-rabbit immunoglobulin as secondary reagent. Antigen-antibody complexes were visualized by chemiluminescence using the ECL Western blotting system (Amersham, Buckinghamshire, United Kingdom).

Chloramphenicol Acetyltransferase (CAT) and Protein Assays—Transfection was performed with 10 μg of reporter plasmid—292 CAT and 1 μg of effector plasmid (cdx2/3 A.cmv, cdx2/3 B.cmv, cdx4.cmv, cmv). RSV-CAT (41) and the cytomegalovirus expression vector were used as positive and negative controls, respectively. Cell extracts were prepared 48 h after transfection and analyzed for CAT activity as described previously (37) except that quantification of acetylated and nonacylated forms was performed using a PhosphorImager. All assays were carried out a minimum of five times. Protein concentrations were determined with a Bio-Rad protein assay kit.

RESULTS

Expression of Homeobox-containing Genes Differs Between Alpha and Beta Cells—To investigate the relative distribution of homeodomain proteins in alpha and beta cells of the endocrine pancreas, we designed degenerate oligonucleotide primers based on highly conserved sequences of the Antennapedia homeodomain class of transcription factors (42). As shown in Table I, cDNA clones corresponding to Idx-1/STF1/IPF1 (11, 16, 43) were most frequently found in insulin-producing cells (HIT-T15 and βTC1) and even the sole sequence obtained from primary islets and primary beta cells. This finding suggests that Idx-1/STF1/IPF1 mRNA levels are the most abundant compared with those encoding other homeodomain proteins of the same class. In glucagon-producing (InR1G9 and αTC1) and primary non-beta cells, by contrast, sequences corresponding to cdx-2/3 were most commonly found. Out of 23 clones analyzed in glucagon-producing cells, 16 matched the sequence of cdx-2 (in mouse-derived cells) or cdx-3 (in hamster-derived cells), whereas 24 out of 28 were found in primary non-beta cells (Table I). Of note, cdx-2/3 sequences were not exclusively found in glucagon- but also in insulin-producing cells. Additional cDNA fragments corresponding to known and unknown homeodomain proteins were amplified in glucagon-producing cells and in primary non-beta cells (Table I). Their respective mRNA levels relative to that of cdx-2/3 were found by RNAse protection analyses to be much lower and for most of them hardly detectable (data not shown).

The relative distribution of homeodomain proteins was analyzed using a PhosphorImager.
Fig. 1. Tissue distribution of the caudal-type homeobox mRNAs. A, schematic diagram representing the location of RNA probes used in RNase protection experiments. cDNAs encoding Cdx-1, Cdx-2/3, and Cdx-4 are illustrated. Shaded boxes represent open reading frames containing...
type homeobox genes, cdx-1, cdx-2/3, and cdx-4 has been reported (for review, see Ref. 50). Since homeobox sequences corresponding to cdx-1 and cdx-4 have also been amplified from rat pancreatic islet cDNA (11), we examined the relative abundance of the three different cdx mRNAs in both glucagon- and insulin-producing cell lines by RNase protection analyses. Using a riboprobe specific for the 5' end of the cdx-1 cDNA (Fig. 1A), a corresponding protected fragment of 412 nt could only be detected in rat intestine, serving as a positive control, but neither in InR1G9 nor in HIT T15 cells (Fig. 1B). These data were confirmed by assays with mRNA derived from mouse aTC1 and rat Ins1 cells. Even after prolonged exposure, no cdx-1-specific signal was visible in these islet cell lines (data not shown).

Similar results were obtained using an antisense in vitro transcript covering the cdx-4 open reading frame. In agreement with the previously reported cdx-4 expression pattern limited to the embryo (23), no cdx-4-specific signal was detected in the adult intestine, InR1G9, or HIT T15 cells (Fig. 1C). Additional assays performed on mouse aTC1, mouse Min6, and rat Ins1 cells also confirmed the absence of cdx-4 transcripts in pancreatic cell lines (data not shown). However, when reverse transcription-PCR amplifications were carried out using cdx-4-specific primers, PCR fragments corresponding to this mRNA could be obtained from InR1G9, primary beta, and primary non-beta cells. Their relative abundance was much lower compared with a control amplification performed on mouse embryonic RNA d9.5 (a kind gift from Dr. J. Huarte, University of Geneva Medical School) (data not shown).

In contrast to cdx-1 and cdx-4 transcripts, which can only be detected by PCR amplification, the cdx-2/3 homeobox gene is expressed at high levels in both insulin- and glucagon-producing cell lines. As shown in Fig. 1D, a fragment of 115 nt corresponding to part of the cdx-2/3 homeobox sequence is protected in InR1G9 and HIT T15 cells as well as in the intestine. The relative abundance of cdx-2/3 transcripts in the pancreatic cell lines tested is similar to that found in the intestine and about 4% that of the b-actin mRNA. We therefore conclude that of the caudal-related genes, only cdx-2/3 is expressed at a significant level in islet cell lines.

Alternative Splicing Generates Two cdx-2/3 mRNAs Present in Islet Cells—Amplification of the cdx-2/3 cDNA using primers 10 and 11 complementary to the 5'- and 3'-ends of the transcript, respectively, generates two specific fragments (Fig. 2A). One PCR product corresponds to the expected size of 1.2 kilobase pairs, whereas an additional smaller band is about 500 bp. Both fragments were identified in glucagon- and insulin-producing cells, in primary islets, and in both primary beta and non-beta cells (data not shown). Sequence analysis of the 500-bp fragment revealed that it shares the 5' and 3' termini with the full-length transcript, whereas it lacks sequences encoding the amino-terminal part of Cdx-2/3. Of note, the homedomain is entirely present in the smaller form, designated Cdx-2/3 B. The translation initiation codon of cdx-2/3 B is found within the 5'-untranslated region of the longer isoform designated cdx-2/3 A, 32 bp upstream of the cdx-2/3 A ATG (Fig. 2A). The sequence upstream of the cdx-2/3 B initiation codon conforms closely to the consensus sequence for translational initiation as proposed by Kozak (51). Interestingly, the 3' splice junction of cdx-2/3 B corresponds to the insertion site of intron 1 of the cdx-2/3 pre-mRNA, whereas the 5' splice junction is located within the ATG start codon of cdx-2/3 A. The 5' and 3' termini of the alternative intron correspond to the consensus splice site sequence of nuclear introns (52) (see Fig. 2A).

To evaluate the relative abundance of the two RNA isoforms in glucagon-producing cells, we performed RNase protection assays on InR1G9 RNA using either the cdx-2/3 A or cdx-2/3 B riboprobe (Fig. 2, B and C). With the cdx-2/3 A antisense in vitro transcript, two major protected bands can be observed, which correspond to the size of the full-length cdx-2/3 A (1068 nt) and the second and third exons (404 nt) comprised in the cdx-2/3 B mRNA, respectively (Fig. 2C). The relative abundance of the truncated mRNA isoform as determined with a PhosphorImager is 30% of all cdx-2/3 transcripts. Similar results were obtained whether total or cytoplasmic RNA was used for the assay (data not shown). Using the cdx-2/3 B riboprobe, completely protected cdx-2/3 B fragments were identified at 534 nt. A second signal at 404 nt, comprising the second and third exon, corresponds in this assay to cdx-2/3 A isoform mRNAs (about 70% of all cdx-2/3 transcripts) and is therefore much more abundant than when using the cdx-2/3 A riboprobe. In addition to the third specific fragment of 130 nt corresponding to part of the first exon, signals that could represent degradation products of the three major protected fragments, splicing intermediates, or alternative mRNA isoforms were detected but at a much lower abundance. We thus conclude that expression of the cdx-2/3 gene in islet cells generates at least two mRNAs by alternative splicing, probably giving rise to two Cdx-2/3 isoforms that differ by their amino-terminal ends. The two cdx-2/3 transcripts are not specific for glucagon-producing cells inasmuch as they are also found in insulin-producing cells and the intestine (data not shown).

To verify that the proteins corresponding to Cdx-2/3 A and B were present in islet cells, we performed Western blot analyses of InR1G9 and HIT T15 nuclear extracts. As shown in Fig. 3A, a band of 34 kDa corresponding to Cdx-2/3 A was observed in both cell types in roughly the same intensity. In contrast to the longer isoform, the anti-Cdx-2/3 antibody was unable to recognize Cdx-2/3 B in Western blot assays as demonstrated by the failure to detect this isoform in clone cdx-2/3 B.S2 overexpressing Cdx-2/3 B. The presence of both Cdx-2/3 proteins in islet cells was confirmed, however, by EMSA with oligonucleotide G1–52 (Table II) matching the consensus binding sequence of caudal-related genes. A complex corresponding to cdx-2/3 A is observed with nuclear extracts from different insulin-, glucagon-, and somatostatin-producing cell lines (Fig. 3B). A second complex representing Cdx-2/3 B is shown in an InR1G9 clone overexpressing the smaller isoform (cdx-2/3 B.S2) and with extracts from RIN-5F and RIN 1027-B2 cells (Fig. 3B). The abundance of Cdx-2/3 B relative to Cdx-2/3 A in both glucagon- and insulin-producing islet cell lines is less than 1:20 on average. This contrasts to the relative levels of both isoform mRNAs (Fig. 2C) and suggests that Cdx-2/3 B has a lower binding affinity for G1–52 or that translational efficiency or degradation rates differ for the Cdx-2/3 isoforms.
matches closely the consensus binding site for caudal-related genes (18, 19). We therefore investigated by EMSA if in vitro produced Cdx-2/3 A and B proteins were able to bind G1. As shown in Fig. 4A, two specific complexes are observed that migrate with the same electrophoretic mobility as the two faster migrating complexes formed with G1–52 and nuclear extracts from islet cell lines or clone cdx-2/3 B.S2. Both complexes are displaced by the addition of anti-Cdx-2/3 antibody and thus correspond to the Cdx-2/3 isoform proteins. The fact that the anti-Cdx-2/3 antibody that was raised against the carboxyl terminus of Cdx-3 (15) recognizes Cdx-2/3 B in EMSA but not in Western blot analysis may possibly be explained by the different protein conformations in both assays. To test for the relative binding affinity of Cdx-2/3 A and B for G1, we competed for both complexes formed with nuclear extracts from clone cdx-2/3 B.S2 with increasing amounts of unlabeled G1–52. As shown in Fig. 4B, competition results in similar displacement effects on Cdx-2/3 A and B complexes, indicating that both isoforms have the same binding affinity for G1; consequently, protein synthesis or degradation rates for Cdx-2/3 A and B must be different. The slowest migrating band was competed with both specific and nonspecific oligonucleotides (data not shown) and does not react with anti-Cdx-2/3 antibody (Fig. 4A).

To analyze in more detail the binding of Cdx-2/3 A to G1, we then investigated protein-DNA complex formation with the oligonucleotide G1–33 containing both AT-rich direct repeat elements of G1 (Table II). As shown in Fig. 4C, incubation of nuclear extracts from BHK 21 cells transfected with the hamster cdx-2/3 A cDNA with oligonucleotide G1–33 results in the formation of two complexes (Cdx-2/3 and D) not present with control extracts from untransfected BHK 21 cells. Both complexes are supershifted by the addition of anti-Cdx-2/3 antibodies but not by preimmune serum. We therefore suggest that the lower and upper complexes represent Cdx-2/3 A monomer and homodimers, respectively. Of note, the monomeric form is much favored over dimer formation in these assay conditions. Using nuclear extracts from glucagon-producing InR1G9 cells, we detect three specific protein complexes, B1, B2, and B3 (Fig. 4C), as previously reported (9). The addition of the anti-Cdx-2/3 antibody displaced B3 but not B1 or B2. Since the migration velocity of B3 is different to both the Cdx-2/3 monomer and dimer (Fig. 4C, BHK 21 extracts), we propose that B3 represents the In1 and In2, for introns 1 and 2, separating the three exons (E1, E2, and E3). A shaded box represents the open reading frame containing the homeobox (black bar). cDNAs cdx-2/3 A and cdx-2/3 B were cloned by PCR amplification with primers 10 and 11, and the splice junction of the shorter form is indicated at the bottom. The ATG start codon of the longer form is underlined and the homeobox sequence is in boldface letters. Lowercase letters indicate sequences of exon 1 that have been deleted in cdx-2/3 B. B, schematic representation of antisense RNA probes (hatched boxes) used for RNase protection analyses and the expected protected fragments (open bars). A dashed line represents the portion of cdx-2/3 mRNA that is not present in cdx-2/3 B; while sequences of the in vitro transcript complementary to vector sequences are indicated by a solid line. C, autoradiogram of RNase protection analyses using riboprobes generated from the hamster cdx-2/3 A and cdx-2/3 B cDNAs. As a positive control, an antisense in vitro transcript (ivt) of the mouse intron insertion site (ivt) was added in all but two assays. The specific activity of this riboprobe was 125 times less than the activity of the cdx-2/3 antisense transcripts. Hybridization was carried out using 50 µg yeast tRNA as a negative control or 30 µg total RNA from the hamster cell line InR1G9. The 130-nt fragment is not visible in lanes 5 and 7 of this autoradiogram because it has a much lower intensity using the cdx-2/3 A riboprobe. This might be due to the fact that the 130-nt fragment represents only 11% of the total length of riboprobe A and 30% of all cdx-2/3 transcripts. ivt, in vitro transcript without RNase treatment. Fragment sizes are given in nucleotides. E, exon; In, intron; UTR, untranslated RNA.
the addition of Cdx-2/3 immune serum (and pancreatic cell lines as indicated with oligonucleotide G1–52 (Table II). Incubation of nuclear extracts from an InR1G9 clone overexpressing Cdx-2/3B (cdx-2/3B.S2) was visualized by chemoluminescence. 

The 34-kDa Cdx-2/3 protein and transferred to a polyvinylidene difluoride membrane for hybridization with anti-Cdx-2/3-specific antibody. The 34-kDa Cdx-2/3 protein was visualized by chemoluminescence.


development with anti-Cdx-2/3 antibody. The 34-kDa Cdx-2/3 protein and transferred to a polyvinylidene difluoride membrane for hybridization with anti-Cdx-2/3-specific antibody. The 34-kDa Cdx-2/3 protein was visualized by chemoluminescence. 


development with anti-Cdx-2/3 antibody. Mutation of the proximal motif by contrast (G1–33r3, Fig. 4D), does not affect B1 or B2, but only formation of B3. Therefore, we conclude that B1 and B2 bind to the distal portion of the G1 control element and do not contain Cdx-2/3 A. In contrast, Cdx-2/3 A binds most preferentially to the proximal AT-rich motif and contributes as a heterodimer to the formation of B3.

We then investigated the effects of more specific mutations within the proximal AT-rich direct repeat motif of G1 on binding of Cdx-2/3 A, using the mutated oligonucleotides G1–52 M11 and M13 (Table II). Both mutations result in a marked decrease of Cdx-2/3 A binding as assayed with InR1G9 and Cdx-2/3 A-containing BHK 21 nuclear extracts (Fig. 4E). To study the effect of the M11 mutation on B3 formation, we used the mutated oligonucleotide G1–56 M11 (Table II); M11 eliminated B3, indicating that the mutated nucleotides are critical for B3 formation (data not shown).

In addition to the two AT-rich sequences arranged as a tandem repeat, G1 contains a third AT-rich motif from nt –54 to –57, that was recently reported to bind the homeodomain factor Isl-1 with high affinity (53). This sequence does not correspond to the consensus binding motif for caudal-related genes and in agreement with this observation, no Cdx-2/3 A complex could be detected with nuclear extracts from InR1G9 cells and an oligonucleotide covering the sequence from –44 to –63 (data not shown). The Cdx-2/3 A binding pattern on G1 observed by EMSAs was also confirmed by DNase I footprint assays. Utilization of Cdx-2/3 A protein overexpressed in BHK 21 resulted in a weak and a strong protection of the distal and proximal AT-rich motif of G1 (data not shown).

Cdx-2/3 A Transactivates the Glucagon Gene Promoter in BHK 21 Cells—We previously reported that point mutations (M11) of the proximal AT-rich direct repeat element within G1 markedly decreased transcriptional activity conferred by the glucagon gene promoter (9). To ascertain whether Cdx-2/3 A binding to this motif leads to changes in transcriptional activity, 292 bp of the 5′-flanking sequence of the rat glucagon gene promoter linked to the CAT reporter gene (–292 CAT) were cotransfected into BHK 21 cells together with increasing amounts of cdx-2/3 A cDNA inserted into a cytomegalovirus promoter-driven expression vector or the vector alone. The reporter construct −292 CAT was inactive in BHK 21 cells when cotransfected with the vector, whereas a dose-dependent

Fig. 3. Relative abundance of Cdx-2/3 proteins in glucagon-, insulin-, and somatostatin-producing cells. A, nuclear extracts of fibroblast cells overexpressing Cdx-2/3 A (5 μg), HIT T15 (30 μg), and InR1G9 cells (30 μg) were separated on a 15% SDS-polyacrylamide gel and transferred to a polyvinylidene difluoride membrane for hybridization with anti-Cdx-2/3-specific antibody. The 34-kDa Cdx-2/3 protein was visualized by chemoluminescence. B, EMSA using 8 μg of nuclear extracts from an InR1G9 clone overexpressing Cdx-2/3 B (cdx-2/3 B.S2) and pancreatic cell lines as indicated with oligonucleotide G1–S2 (Table II). The addition of Cdx-2/3 immune serum (I) is indicated. P" designates free oligonucleotides, and nonspecific complexes are indicated by an asterisk.

Table II

| Synthetic oligonucleotides used in EMSA |
|----------------------------------------|
| Mutations in oligonucleotides deviating from the wild-type rat glucagon gene sequence are indicated. Nucleotides in the G1 element homologous to binding sites of caudal-related homeodomain proteins (18) are underlined. |
| G1–56 | AGTGAATCATTTGACAAAAACCCCCATTATTACAGATGGAAGATTATATTGTTCAG |
| G1–56 M11 | -----------------------------------------------CC------------- |
| G1–52 | GAGAAATTTATTTGCAGCG |
| G1–52 M11 | -----------------------------------------------CC------------- |
| G1–52 M13 | -----------------------------------------------CC------------- |
| G1–33 | CCCCCATTTATTTGAGAAATTTAATTGT |
| G1–33r5 | ---------------------CCCCCAG------------------- |
| G1–33r3 | ---------------------GACAGCCCCGC------------------- |
activation in transcription resulted from overexpression of Cdx-2/3 A (Fig. 5).

Interference of Cdx-2/3 B with Cdx-2/3 A Binding on G1 and Transcriptional Activation—To analyze the impact of Cdx-2/3 B on DNA-protein complexes formed on the G1 element, we stably transfected InR1G9 cells with the smaller isoform cloned

![Graph showing binding of Cdx-2/3 to the proximal promoter of the glucagon gene G1.](image)
in the eukaryotic expression vector pBJ1-neo (34). When nu-
clear extracts of the generated InR1G9 clones are incubated
with oligonucleotide G1–52, different levels of Cdx-2/3 B over-
expression can be observed (Fig. 6A). Whereas clone B.S7 does
not express the truncated isoform, clones S.B2 and S.B5 con-
tain an equal level and twice the level, respectively, of Cdx-2/3
B relative to Cdx-2/3 A. Incubation of the same nuclear extracts
with G1–56 containing both AT-rich sequences results in the
formation of the usual complexes B1, B2, and B3 and of a
complex representing the Cdx-2/3 B monomer (Figs. 4C and
6B). Of note, the intensity of B3 is decreased with nuclear
extracts from all clones overexpressing Cdx-2/3 B, although
there appears to be no strict correlation with the expression
level of the truncated isoform. Additionally, a new complex is
detected from clone B.S5, probably representing a homo-
or heterodimer of Cdx-2/3 B. Of note, the new complex is only
formed with InR1G9 extracts expressing high levels of Cdx-2/3
B, indicating that binding of the monomer is favored over dimer
formation, an observation reinforced by the failure to detect
Cdx-2/3 B homodimers after overexpression in BHK 21 cells
or protein synthesis in vitro (data not shown, Fig. 6C). To inves-
tigate complexes formed on G1 in the presence of both isoforms
overexpressed in BHK 21, we mixed the two nuclear extracts
prior to the addition of G1–56. As seen in Fig. 6C, no additional complexes that might contain Cdx-2/3 A and Cdx-2/3 B could be observed. Our data indicate that Cdx-
2/3 A or B homodimer and Cdx-2/3 A/B heterodimer formation
are not readily observed on G1 and that heterodimers between
Cdx-2/3 A or B and another protein are probably favored.

To compare the effects of both Cdx-2/3 isoforms on transcriptional
activation, we cotransfected either cdx-2/3 A or cdx-2/3 B with −292 CAT. While cotransfection with cdx-2/3 A re-
resulted in a 11-fold increase, cotransfection of cdx-2/3 B was, by
contrast, without any effect (Fig. 7A). The fact that Cdx-2/3 B is
able to bind the glucagon promoter without activating tran-
scription suggests that the activation domain of Cdx-2/3 might
be located in the N-terminal part of the protein, which is
deleted in isoform Cdx-2/3 B. To investigate whether the trun-

---

**Fig. 5.** Dose-response curve of transactivation of the glucagon promoter by Cdx-2/3. Different amounts of cdx-2/3 A cDNA in a
cytomegalovirus-driven expression vector (0.1, 0.5, and 1 µg of plasmid DNA) (top line) were cotransfected into BHK 21 cells together with a
reporter plasmid containing the CAT gene under the control of sequence nt −292 to +58 of the rat glucagon gene promoter (−292 CAT; Ref. 7).
Cotransfection of the expression vector alone (bottom line) served as a
negative control. CAT activities assayed after 48 h are given relative to
the positive control RSV-CAT as the mean of eight experiments ± S.E.

**Fig. 6.** Interference of Cdx-2/3 B with complex formation on
G1. A and B, EMSAs using nuclear extracts from InR1G9 clones stably
transfected with cdx-2/3 B and oligonucleotides G1–52 and G1–56. B1, B2,
and B3 indicate the positions of specific complexes formed on G1
B1 and G2. * designates free oligonucleotides, and nonspecific complexes are indicated by an asterisk.
of Cdx-2/3 A compared with Cdx-2/3 B as assayed by EMSAs (data not shown). To further analyze the effect of Cdx-2/3 B expression in glucagon-producing cells, we transfected InR1G9 cells producing only little amounts of the smaller isoform and also the two stably transfected clones cdx-2/3 B.S2 and cdx-2/3 B.S5 with the reporter construct −292 CAT. As shown in Fig. 7B, expression of relatively higher levels of Cdx-2/3 B results in a 72% reduction of relative transcriptional activity compared with InR1G9 cells. In contrast, no significant decrease in activity is observed when similar levels of the Cdx-2/3 isoforms are present in InR1G9 cells. The presence of Cdx-2/3 B in a higher concentration relative to Cdx-2/3 A may interfere with complex formation on G1 by competing for the protein partners of Cdx-2/3 A and therefore influence the relative transcriptional activation by a mechanism that remains to be determined.

**DISCUSSION**

Differentiation of multicellular organisms is based on a precise temporal and spatial pattern of cell specific gene expression. In vertebrates, as in all other metazoa analyzed today, homeotic genes are major control genes determining cell fate and function (42). Islet cells express a variety of homeobox genes, and recent data indicate that Idx-1/STF1/IPF1, a homeodomain protein of the Antennapedia class, plays a major role in the cell-specific expression of the insulin gene, its regulation by glucose, and the differentiation of the pancreas (16, 43, 54–57).

To better understand the molecular mechanisms involved in the islet cell specific expression of the glucagon gene, we searched for homeotic cDNA sequences in glucagon-producing cells by reverse transcription-PCR. cdx-2/3 appears from our PCR analysis to be the predominant homeobox gene from the Antennapedia class in glucagon-producing cells, although it is also present in the intestine and in pancreatic beta cells (15, 17, 18). Quantification of cdx-2/3 by RNase protection experiments, Western blots, and EMSAs on RNA and nuclear extracts from glucagon-, insulin-, and somatostatin-producing cell lines show roughly equivalent cdx-2/3 transcript and protein levels in all cell types. In addition, the amount of PCR-amplified full-length cdx-2/3 cDNA products from primary beta and non-beta cells does not indicate major differences. The reason for the preferential amplification of Idx-1/STF1/IPF1 sequences in primary beta cells is then likely due to the relatively higher Idx-1/STF1/IPF1 mRNA levels compared with cdx-2/3. The relative distribution of Antennapedia class homeobox sequences identified in our study is slightly different from previous PCR analyses in islet cell lines and primary islets (11, 12); some homeobox genes reported to be expressed in primary islets could not be detected among the cDNA clones sequenced (Table I). Likely explanations for these differences are statistical variance, primer choice, and the source of RNA analyzed.

Cdx-2/3 is a member of the family of homeodomain proteins related to the Drosophila melanogaster gene caudal (cad) (58). caudal-related genes have been found in Bombyx mori (59), Caenorhabditis elegans (60), zebrafish, Xenopus laevis, and chicken, as well as in the mammals mouse, rat, hamster, and human (50, 61). In vertebrates, expression onset of most caudal-related genes is observed at the state of gastrulation exhibiting a gradient along the anteroposterior axis. With the exception of Chox-cad2 (62) and Xcad-3 (63), transcripts of caudal-related genes are detected predominantly in endodermally derived tissues, playing a major role in the development of the intestine (17, 23, 64, 65). For cdxA, early expression pattern is reported to cover several organs of endodermal origin like liver, pancreas, lung, and the epithelial lining of the intes-
Glucagon Gene Expression

U.S. Food and Drug Administration

Glucagon Gene Expression

28993

tine, whereas later expression is restricted to the intestine. The presence of cdx-2/3 in adult cells of the endocrine pancreas (15, 66) may indicate a wider role of cdx genes not only in the development of endodermally derived tissues but also in transcriptional control of differentiated islet cells. In contrast to the intestine, cdx-2/3 is the only caudal-related gene expressed at significant levels in the islets of Langerhans. Interestingly, alternative splicing of the cdx-2/3 mRNA leads to an amino-terminally truncated isoform protein with translational initiation at an upstream AUG normally supposed to be skipped (15). Of note, all mammalian caudal-related genes exhibit a second AUG codon 22–33 nt upstream of the initiator AUG (17, 23, 36). With the exception of cdx-2/3 B, however, these upstream AUGs only imperfectly match the Kozak consensus for vertebrates, which may explain the deviation of the first AUG rule. The initiator context for cdx-2/3 B (GGAGCAUGG) is close to the consensus sequence GCCACCAUGG, allowing translational initiation in the absence of the second AUG. Protein synthesis of Cdx-2/3 B may not be optimal, however, inasmuch it is much lower than synthesis of Cdx-2/3 A in in vitro transcription/translation assays.

Our results indicate that one of the functions of cdx-2/3 may be the regulation of glucagon gene expression. Cdx-2/3 indeed binds the upstream control element G1 at the proximal AT-rich repeat motif, which corresponds to the consensus binding site for caudal-related genes (19). It is also capable, although with much lower affinity, of binding the distal, imperfect repeat motif within G1. However, in contrast to the proximal motif, the distal element contains a C at position 6, where in the CdxA consensus as well as in the Cad binding sites there is a strong preference for T. When oligonucleotides used for EMSAs contain both repeat elements, Cdx-2/3 A homodimer formation (complex D) is observed with BHK 21 nuclear extracts containing high concentrations of Cdx-2/3 A and disappears when one of the repeat elements is mutated. Our data indicate that if homodimer formation occurs, this is only in the presence of large amounts of Cdx-2/3 A. With cdx-2/3 A-transfected BHK 21 extracts the intensity of complex D is less than 1⁄10 that of the monomeric complex. Cdx-2/3 A homodimer formation is thus not favored with G1, an observation in agreement with the low affinity of the distal AT-rich motif for Cdx-2/3 A. By contrast, with InR1G9 nuclear extracts, only the slow migrating complex B3 is recognized by anti-Cdx-2/3 antibodies, and a Cdx-2/3 A monomer can only be detected using small or mutated oligonucleotides that may prevent the assembly of higher molecular weight complexes. Thus, with extracts from glucagon-producing cells, formation of heterodimers is much favored over monomers. Caudal binding sites in Drosophila (67), the rat insulin I gene (15), and the mouse sucrase-isomaltase gene (18) represent direct or indirect repeats, and for the latter homodimer formation was observed. Although most dimer-forming homeodomain proteins bind to sequences with dyad symmetry, it is well established that dimerization also occurs on tandem repeats such as G1 (68–71). However, in contrast to the redox potential-dependent dimer formation on the sucrase-isomaltase gene promoter, neither the Cdx-2/3 A homodimer nor B3 binding to G1 is influenced by the redox state (data not shown). In a recent report, Jin and Drucker (66) have observed binding of Cdx-2/3 A as a monomer to the G1 element. However, the oligonucleotides used for EMSA in their study comprised only one Cdx-2/3 binding site and thus do not enable dimer formation.

Binding of Cdx-2/3 A to the glucagon promoter leads to a dose-dependent activation of transcription in non-glucagon-producing cells, and mutation of the Cdx-2/3 A binding site leads to a marked decrease in transcriptional activity, suggesting a functional relevance for Cdx-2/3 A in glucagon gene expression (9). The presence of an amino-terminally truncated Cdx-2/3 isoform in glucagon-producing cells that binds to G1 but does not activate transcription opens additional possibilities for gene regulation. The interference of Cdx-2/3 B with the transcriptional property of Cdx-2/3 A in BHK 21 and InR1G9 cells suggests a role of Cdx-2/3 B as a modulator of glucagon gene expression. This effect could be due to competition of both isoforms for binding to the glucagon promoter or to protein-protein interactions inhibiting the activation properties of Cdx-2/3 A. It will be interesting to analyze if the alternative splicing leading to both Cdx-2/3 isoforms is differentially regulated during development or in response to physiological stimuli and correlated to changes in glucagon gene expression.

G1 is a large DNA control element that likely interacts with multiple transcription factors; Cdx-2/3 A may be one of them present in complex B3. We previously identified at least three additional protein complexes, B1, B2, and B6 (9), that might also be functionally important for glucagon gene expression. Recently, the LIM-domain homeobox gene ins-1 was shown to interact with oligonucleotides containing the most proximal AT-rich motif (positions 54 to 57) of the glucagon promoter element G1 (53). However, in cotransfection experiments with 292 bp of the 5'-flanking sequence of the rat glucagon gene promoter linked to the CAT reporter gene, ins-1 was unable to activate transcription, and no synergistic effect with Cdx-2/3 A on transcription of the glucagon gene could be observed. Therefore, Cdx-2/3 A is unlikely to interact with ins-1, which by itself does not appear to play a major role in glucagon gene transcription.

Cdx-2/3 A has been implicated in the control of cell proliferation and differentiation in the intestine (65) as well as in the regulation of intestine-specific gene transcription, at least for the sucrase-isomaltase, carbonic anhydrase I, and lactase genes (18, 72). Cdx-2/3 A is also able to bind and activate the rat insulin I gene promoter at the minienhancer (15), although other homeodomain factors such as Idx-1/STF1/IPF1 are more likely physiological activators of insulin gene transcription (11, 16, 54–56). Our data and a recent report (66) indicate that Cdx-2/3 A is involved in glucagon gene expression both in the endocrine pancreas and in enteroendocrine cell lines expressing glucagon. In both cell types, Cdx-2/3 A was shown to transactivate the glucagon gene promoter (66). These results might be consistent with the endodermal origin of both tissues and indicate that some of the mechanisms of glucagon gene regulation in the intestine and endocrine pancreas might be similar. The presence of two Cdx-2/3 isoforms with different transcription properties not only in glucagon- but also in insulin-producing cells as well as in the intestine opens further possibilities to control gene expression in these tissues.

Acknowledgment—We thank B. Mach and his group for helpful advice.

REFERENCES

1. Drucker, D. J., and Asa, S. (1988) J. Biol. Chem. 263, 13475–13478

2. Mojsov, S., Heinrich, G., Wilson, I. B., Ravazzola, M., Orci, L., and Habener, J. F. (1986) J. Biol. Chem. 261, 11880–11889

3. Novak, U., Wilks, A., Buellin, G., and McEwen, S. (1987) Eur. J. Biochem. 164, 553–558

4. Efrat, S., Teitelman, G., Anwar, M., Ruggiero, D., and Hanahan, D. (1988) Neuron 1, 605–613

5. Lee, Y. C., Asa, S. L., and Drucker, D. J. (1992) J. Biol. Chem. 267, 10705–10708

6. Cordier-Bussat, M., Morel, C., and Philippe, J. (1995) Mol. Cell. Biol. 15, 3904–3916

7. Philippe, J., Drucker, D., Knevel, W., Jepal, L., Misulovin, Z., and Habener, J.
