THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1999 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 274, No. 9, Issue of February 26, pp. 6011–6019, 1999
Printed in U.S.A.

Identification of Domains Mediating Transcriptional Activation and Cytoplasmic Export in the Caudal Homeobox Protein Cdx-3*

(Received for publication, November 20, 1998, and in revised form, December 3, 1998)

Khiet Y. Trinh‡, Tianru Jin§, and Daniel J. Drucker¶

From the Department of Medicine, The Toronto Hospital, Banting and Best Diabetes Centre, University of Toronto, Toronto, Ontario M5G 2C4, Canada

The *caudal* genes have important functions in embryonic development and cell differentiation. The *caudal*-related protein Cdx-2/3 (the protein designated Cdx-2 in the mouse and Cdx-3 in the hamster) is expressed in the gastrointestinal epithelium and in islet and enteroendocrine cells, where it activates proglucagon gene transcription. We show here that Cdx-3 sequences amino-terminal to the homeodomain (amino acids 1–180) function as a heterologous transcriptional activation domain when fused to the LexA DNA binding domain. A Cdx-3-Pit-1 fusion protein containing only the first 83 amino acids of Cdx-3 linked to the POU domain of Pit-1 markedly stimulated the transcriptional activity of a Pit-1-responsive promoter. Analysis of the transcriptional properties of Cdx-3 mutants in fibroblasts and islet cells revealed distinct amino-terminal subdomains that function in a cell-specific manner. Point mutations within the amino-terminal A domain were associated with reduced transcriptional activity. Furthermore, internal deletions and selected point mutations within domain A, but not the B or C domains, resulted in accumulation of mutant Cdx-3 in the cytoplasm. Unexpectedly, mutation of an Asp-Lys-Asp motif within domain A identified a putative cytoplasmic membrane-associated export signal that mediates Cdx-3 compartmentalization. These experiments delineate unique activities for specific amino-terminal sequences that are functionally important for Cdx-3 biological activity.

The *caudal* genes, originally identified in *Drosophila* (1), are expressed early during development in a concentration gradient across the anteroposterior axis of the embryo. Perturbation of *caudal* function as a result of inactivating mutations (2) or ectopic expression (3) leads to disruption of normal segmentation and an abnormal body pattern. *Caudal* genes were subsequently identified and characterized in mammalian species, including mice and humans (4–6). The mammalian *caudal* homologues *cdx*-1 and *cdx*-2/3 are expressed early during embryonic development in the gut, tailbud, neural tube, and mesoderm, following which, expression during adult life is predominantly restricted to the epithelium of the gastointestinal tract (4, 7). In contrast, the *cdx*-4 gene, which maps to the X-inactivation center of mouse and human X chromosomes (8), is expressed only transiently during embryonic development, from day 7 to day 10 post coitum (9).

Several lines of evidence suggest that members of the *caudal* gene family have important roles in the control of cell proliferation and differentiation in cells of endodermal origin. The chicken Caudal protein CHox-cad is expressed in embryonic but not adult liver, and expression is reactivated in regenerating liver cells following partial hepatectomy, consistent with a role for CHox-Cad in hepatocyte differentiation (10). Conditional expression of murine *cdx*-2 in undifferentiated intestinal IEC-6 cells leads to cessation of cellular proliferation followed by resumption of growth and development of cellular characteristics reflective of more differentiated intestinal cell types (11). Overexpression of *cdx*-2 in human colon cancer cells up-regulates molecules involved in cell adhesion and cell-cell interactions, and *cdx*-2 expression in turn is modulated by basement membrane components (12). Taken together, these experiments provide substantial evidence linking *cdx*-2 with changes in cell growth and differentiation.

Complementary evidence for the biological importance of *caudal* genes in pattern formation and intestinal differentiation derives from analysis of loss of function mutations in mice. Disruption of the murine *cdx*-1 gene by homologous recombination results in anterior homeotic transformation of vertebrae, potentially due to the importance of *cdx*-1 for transcriptional control of members of the *Hox* gene cluster (13). Although homozygous *cdx*-2−/− mice are embryonic lethal, *cdx*-2+/− heterozygote mice exhibit a complex phenotype, with tail and skeletal abnormalities including a homeotic shift of vertebrae and rib malformations (14). Cdx-2+/− mice also exhibit stunted growth and develop multiple adenomatous intestinal polyps that do not express the remaining *cdx*-2 allele. These observations demonstrate that even *cdx*-2 haploinsufficiency leads to striking abnormalities in cellular differentiation pathways *in vivo*.

Despite the emerging biological importance of Caudal proteins, the structural determinants of *caudal* function remain poorly understood. Alignment of *caudal* family members demonstrates significant amino acid homology within subdomains, designated A–C (9, 15) of the amino-terminal region; however, the functional significance of these subdomains has not yet been ascertained. We recently demonstrated that the hamster *cdx*-2 homologue, designated *cdx*-3, is an important regulator of proglucagon gene transcription not only in pancreatic islets but also in enteroendocrine cells of the intestine (16, 17). To elucidate the specific regions of the Cdx-2/3 protein (the protein designated Cdx-2 in the mouse and Cdx-3 in the hamster) that are important for transcriptional activation and biological function, we have now carried out studies of the structural deter-

* This work was supported in part by an operating grant from the Medical Research Council of Canada. 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.

‡ Banting and Best Diabetes Center Research Fellow and recipient of a Medical Research Council of Canada Clinician Scientist Award.

§ Supported by a Research Fellowship from the Ontario Ministry of Health

¶ Scientist of the Medical Research Council of Canada. To whom correspondence and reprint requests should be addressed: The Toronto Hospital, 200 Elizabeth St., CCRW3-838, Toronto M5G 2C4, Ontario, Canada. Tel.: 416-340-4125; Fax: 416-978-4108; E-mail: d.drucker@utoronto.ca.

minants of Cdx-3-dependent activation of the proglucagon promoter *in vitro*.

MATERIALS AND METHODS

Plasmids—The plasmid pBAT7/8 and *cdx*-3 were kindly provided by Dr. M. S. German (San Francisco, CA). The plasmid pBAT8 is similar to pBAT7 (18), except that additional restriction sites were introduced into the polylinker to facilitate cloning. The pBAT8 expression vector is under the control of the cytomegalovirus promoter. The linker 5′-CCTAGGACACCATGTACGTGAGCTACAAGCTTCTCGAG-3′, containing an optimized Kozak consensus sequence and encoding the first seven amino acids of *cdx*-3, was ligated into pBAT8 at the Avr II and *Xho*I sites. To construct *cdx*-3 deletions, *Hin*dIII and *Xho*I restriction sites were introduced into the linker sequences, resulting in the insertion of a lysine at position 6 after tyrosine at position 5 and a change of aspartic acid to glutamic acid at position 8 of the Cdx-3 amino terminus. The polymerase chain reaction (PCR)[1] was used to synthesize various amino-terminal, carboxyl-terminal, and internal deletions of *cdx*-3. The *cdx*-3 mutant cDNAs were cloned into the *Xho*I site of the expression vector pBAT8. The open reading frame for all mutant *cdx*-3 cDNAs was verified by DNA sequencing to ensure that the DNA and corresponding translated protein sequence was correct and in frame.

The full-length Pit-1 cDNA in the expression vector Px6.3 and the LexA/luciferase reporter and LexA-Pit-1(8–80) plasmids were obtained from Dr. H. Elsholtz (University of Toronto) and have been described (19, 20). The chimeric LexA-*cdx*-3 cDNAs containing various lengths of *cdx*-3 coding sequences were generated by fusing *cdx*-3 sequences to a cDNA encoding an 87 amino acid LexA DNA binding domain. The −82 proglucagon promoter-luciferase reporter gene has been previously described (16). Site-directed mutagenesis was carried out using the Amersham Pharmacia Biotech unique site elimination mutagenesis kit, optimized for PCR (21). All mutant *cdx*-3 cDNAs were subject to DNA sequencing to verify the presence of the correct sequence change.

Cell Culture and Transfections—Baby hamster kidney (BHK) fibroblasts and InR1-G9 islet cells (22) were grown in Dulbecco's modified Eagle's medium supplemented with 5% calf serum. Cells were transfected using the calcium phosphate method and harvested 18 h after transfection for determination of luciferase activity using a Berthold LB9501 luminometer. All transfections were carried out in triplicate on at least three separate occasions, with 5 μg each of luciferase reporter plasmid and the *cdx-3* or LexA-*cdx* cDNAs.

Antiserum to cdx-3 and Western Blot Analysis—A hamster *cdx*-3 partial cDNA generated by PCR that encodes the carboxyl-terminal 54 amino acids was ligated in frame downstream of glutathione *S*-transferase sequences in a pGEX plasmid (Amersham Pharmacia Biotech). The glutathione *S*-transferase-Cdx-3 fusion protein was purified using glutathione-Sepharose beads and used for antibody production in rabbits. For Western analyses, 5–10 μg of nuclear protein was size-fractionated by SDS-PAGE and transferred to a nitrocellulose membrane. The immunoreactive protein was detected by the ECL system (Amersham Pharmacia Biotech) with 1:2000 dilution of the primary Cdx-3 antiserum and 1:5000 dilution of the peroxidase-labeled anti-rabbit IgG as second antibody. Polyclonal c-Jun antibody (1:1000) was from Sant Cruz Biotechnology, Inc. (Santa Cruz, CA).

Nuclear Protein Extract and EMSA—Nuclear protein was isolated from transfected cells as described previously (16). Synthetic oligonucleotides corresponding to the proglucagon proximal promoter G1 sequences (16, 23) were annealed, purified by 5% polyacrylamide gel electrophoresis and ^{32}P-labeled with the Klenow enzyme. EMSA was performed as described previously (16, 17). A total of 40,000 cpm of ^{32}P-labeled double-stranded oligonucleotide probe was used in each EMSA reaction.

Fractionation of Nuclear and Cytoplasmic Membranes—Nuclear and cytoplasmic fractionation was carried out essentially as described (24, 25). Transfected BHK cells were harvested in ice-cold phosphate-buffered saline, washed twice in the same buffer, and then resuspended in 500 μl of low salt-detergent buffer containing 10 mM Tris, pH 7.8, 1% Nonidet P-40, 10 mM β-mercaptoethanol, 0.5 mM phenylmethylsulfonyl fluoride, 10 μg/ml trypsin inhibitor, and 1 μg/ml aprotinin at 4 °C for 5 min. The swollen cells were lysed by 10 passages through a 22 gauge needle. Crude nuclei were pelleted at 1000 × *g* for 5 min at 4 °C, further purified by resuspension in 2.2 M sucrose, and centrifuged at 80,000 × *g* for 90 min. The supernatant of the low-speed spin was immediately centrifuged at 100,000 × *g* for 30 min at 4 °C to obtain the cytoplasmic plasma membrane fraction in the pellet. The purified nuclei were digested with 250 μg/ml of RNase A and DNase I at 4 °C for 60 min and incubated with high salt buffer containing 2 M NaCl, 10 mM Tris, pH 7.5, 0.2 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride, and 1% (v/v) β-mercaptoethanol at 4 °C for 15 min to rupture the nuclei, then dialyzed against phosphate-buffered saline buffer. An aliquot of the disrupted nuclei was removed for analysis as nuclear extract. The broken nuclei were spun at 1600 × *g* for 30 min at 4 °C, and the pellet containing nuclear membranes was resuspended in 20% glycerol, 1 mM EDTA, and 10 mM Tris, pH 7.5, designated the nuclear membrane fraction, and stored at −20 °C.

RESULTS

The results of recent experiments suggest that amino acids 1–180, which make up the amino-terminal region of *cdx*-2, function as a heterologous activation domain (26). To identify functionally important regions of the Cdx-3 protein independent of the Cdx-*3* homeobox, we generated a series of LexA-Cdx-3 fusion proteins. The transcriptional properties of these molecules were studied using a cotransfected reporter plasmid containing LexA DNA binding sites adjacent to a minimal promoter upstream of the luciferase cDNA coding sequences. As a control, we first analyzed the transcriptional properties of a LexA:Pit-1 cDNA containing the LexA DNA binding domain fused to amino acids 8–80 of the transcription factor Pit-1 (19). The Pit-1 amino-terminal sequences conferred approximately a 3-fold induction of reporter activity in transfected BHK cells, similar to the magnitude of induction reported with the identical LexA:Pit-1 plasmid in transfected COS-7 cells (19). Fusion of the entire *cdx*-3 open reading frame adjacent to the LexA DNA binding domain resulted in a 6-fold activation of reporter activity (Fig. 1*a*).

Although deletion of the sequences carboxyl-terminal to the *cdx*-3 homeodomain (LexA:*cdx* 1–250) was associated with a reduction in relative transcriptional activation, the *cdx*-3 carboxyl-terminal sequences alone, from amino acids 245–313, did not confer transcriptional activation to the LexA DNA binding domain. The greatest activation of luciferase activity was observed with the plasmid containing the *cdx-3* amino-terminal sequences, LexA:*Cdx*(1–180), resulting in a 25–40-fold activation of reporter activity (Fig. 1, *a* and *b*). Plasmids containing the homeodomain alone, or the homeodomain plus the carboxyl-terminal sequences, displayed no transcriptional activation (Fig. 1*a*).

The results of these experiments using chimeric proteins in transfected fibroblasts demonstrated that Cdx-3 amino acids 1–180 encode a transcriptional activation domain, in agreement with experiments carried out analyzing the transcriptional properties of the *cdx*-2 amino-terminal region (26). To determine whether protein subdomains within the 180 amino acid Cdx-3 amino-terminal region might be identified that are sufficient for conferring transcriptional activation, we analyzed the activity of a series of LexA:*cdx* fusions containing smaller regions of Cdx-3 amino-terminal sequences. In contrast to the potent transcriptional activation observed with LexA plasmids containing the intact 180 amino acid amino-terminal region of Cdx-3, LexA:*cdx* fusion genes containing smaller regions of *cdx*-3 amino-terminal sequences were transcriptionally inactive (Fig. 1*b*).

These experiments demonstrated that Cdx-3 amino-terminal sequences display functional activity when fused to the DNA binding domain of a bacterial transcription factor. To determine whether Cdx-3 amino-terminal sequences also mediate transcriptional activation in the context of a heterologous mammalian transcription factor, the first 120 amino acids of Pit-1 were replaced with amino acids 1–83 of Cdx-3. Transfection of BHK fibroblasts with the Cdx-3(1–83)/Pit-1 chimeric protein produced a 350-fold activation of luciferase activity

---

[1] The abbreviations used are: PCR, polymerase chain reaction; EMSA, electrophoretic mobility shift assay; BHK, baby hamster kidney.

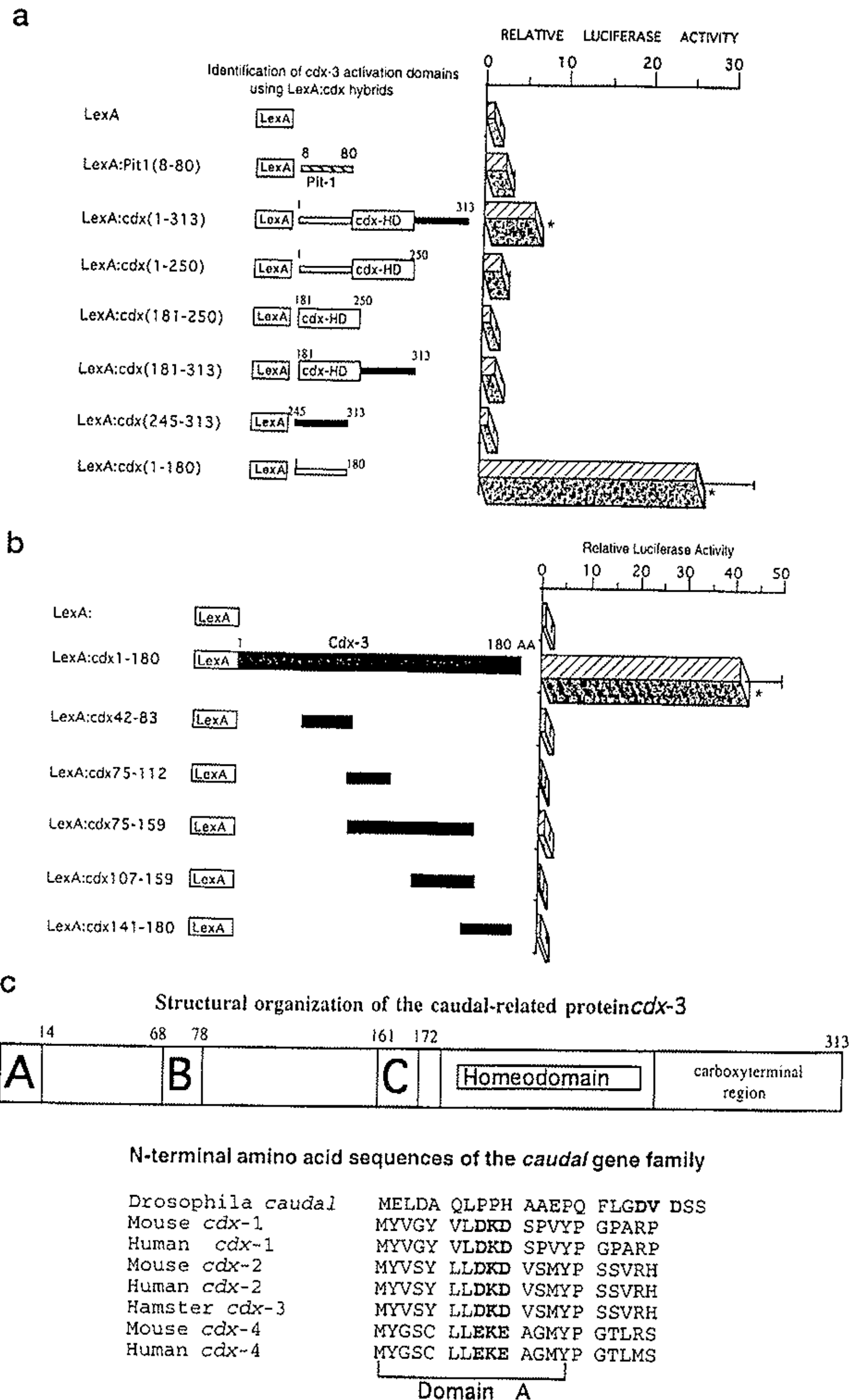

FIG. 1. *a* and *b*, transcriptional properties of the LexA:Cdx-3 chimeric proteins in BHK fibroblasts. Cdx-3 or Pit-1 sequences generated by PCR were fused to the LexA DNA binding domain. Each LexA plasmid (5 μg) was cotransfected with a reporter plasmid containing the LexA operator site fused to the luciferase coding sequences. The relative luciferase activity was normalized to the values obtained following transfection of the LexA vector containing only the DNA binding domain and the luciferase reporter alone. All PCR products were sequenced to verify the absence of PCR-induced mutations. Each LexA plasmid was transfected in triplicate, and the data shown represent the mean ± S.E. of three separate experiments. *, $p < 0.05$ (experimental compared with transfection with control plasmid with the LexA DNA binding domain alone). *c*, structure of hamster Cdx-3 and conserved subdomains within the amino-terminal region of the Caudal-related proteins. Domains A (amino acids 1–14), B (amino acids 68–78), and C (amino acids 161–172) within *cdx*-3 are shown. The amino-terminal amino acid sequences from members of the caudal family are shown. The putative endoplasmic reticulum export signal (D/E)*X*(E/D), located within domain A, is shown in *boldface*. Protein sequences were from GenBank™ (accession numbers 157045 (Drosophila *caudal*), 1170313 (mouse *cdx*-1),1345720 (human *cdx*-1), 1170314 (mouse *cdx*-2), 2495272 (human *cdx*-2), 423360 (hamster *cdx*-3), 729725 (mouse *cdx*-4), and 3023467 (human *cdx*-4)).

(Fig. 2), significantly greater than the 100-fold activation of reporter activity observed following transfection of full-length wild-type Pit-1 (Fig. 2). In contrast, minimal activation was seen with the control plasmid Pit-1(121–191). These observations define a subregion of the Cdx-3 amino terminus (amino acids 1–83, containing subdomains A and B (Fig. 1*c*)) as being capable of functionally substituting for an endogenous activation domain in a mammalian transcription factor.

As the above experiments used heterologous DNA binding domains and DNA target sequences for analysis of Cdx-3 transcriptional activation, we next examined whether the amino-terminal Cdx-3 sequences were functionally important for transcriptional activation of the proglucagon promoter, a known *cdx*-3-responsive target gene (16, 27). A proglucagon promoter-luciferase reporter plasmid was cotransfected with a series of *cdx*-3 plasmids containing in-frame deletions in the *cdx*-3 amino-terminal coding sequence. Full-length *cdx*-3 produced an ~8–17-fold activation of the proglucagon promoter (Fig. 3, *a* and *b*). Remarkably, internal deletion of Cdx-3 amino acids 84–180, containing sequences comprising subdomain C (Fig. 1*c*), had no effect on the relative transcriptional induction of the proglucagon promoter (Figs. 3*a* and 4). In contrast, deletions within the first 75 amino acids containing domains A and B, alone or in combination with more extensive carboxyl-

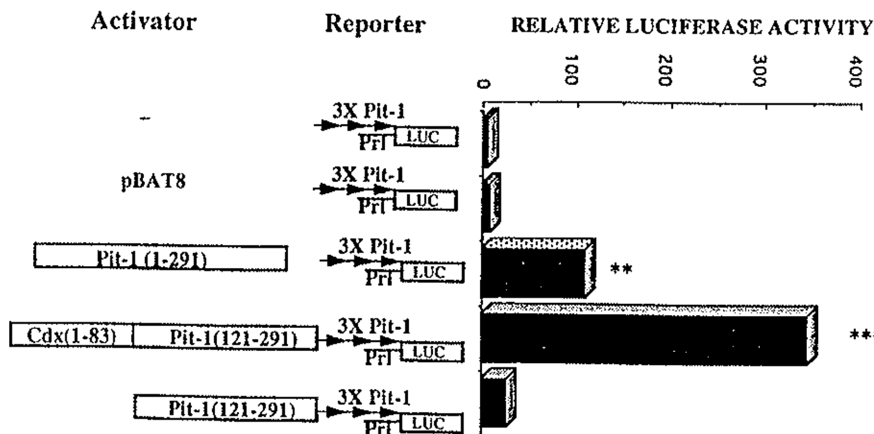

FIG. 2. **Transactivation by a chimeric Cdx-3/Pit-1 protein.** Relative transcriptional activation by full-length Pit-1 *versus* a chimeric Cdx-3/Pit-1 protein following transfection of BHK cells with a Pit-1-responsive luciferase reporter plasmid. The relative luciferase activity obtained following transfection of the Pit-1-responsive luciferase reporter alone, designated −, with the cytomegalovirus expression vector alone (pBAT8) or with the Pit-1(1–291), Pit-1(121–291), or *cdx*-3/Pit-1 chimeric expression vectors is shown. The relative luciferase activity was normalized to values obtained following transfection of the luciferase reporter alone. Each plasmid was transfected in triplicate, and the data shown represent the mean ± S.E. of three separate experiments. **, $p < 0.01$; ***, $p < 0.001$ (experimental compared with control Pit-1-responsive luciferase plasmid alone).

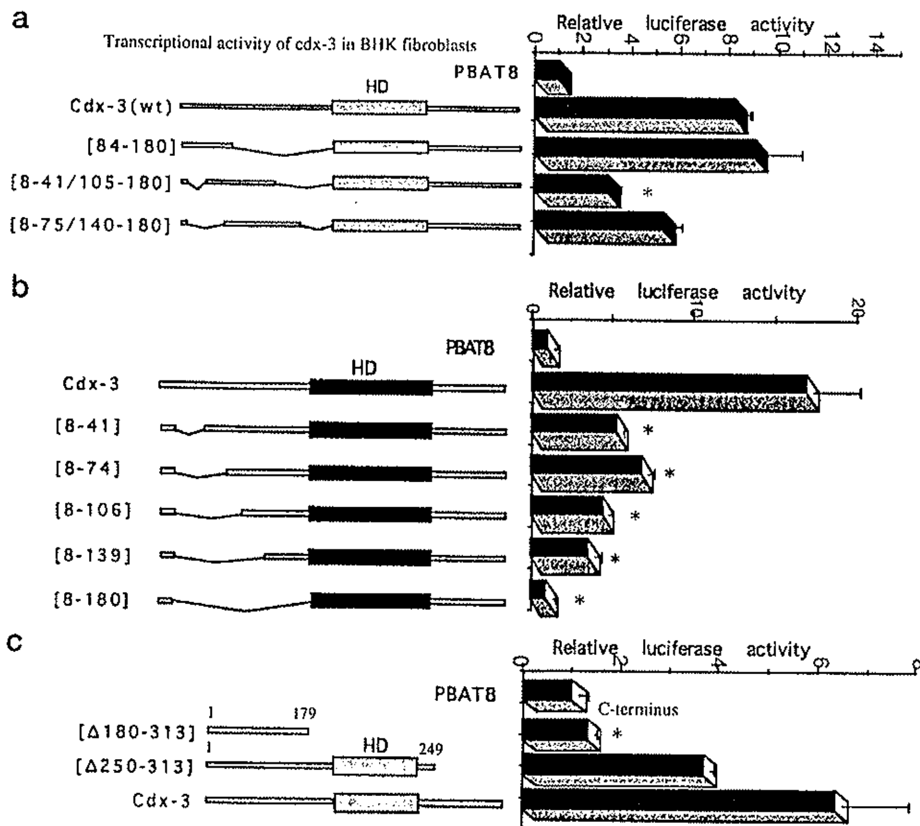

FIG. 3. **Transcriptional properties of *cdx*-3 mutant plasmids.** A series of amino-terminal deletions (*a* and *b*) or carboxyl-terminal deletions (*c*) was generated by PCR. The *cdx*-3 plasmids (5 μg) were cotransfected into BHK cells with 5 μg of the −82 proglucagon-luciferase plasmid, and the luciferase activity obtained was normalized to the background luciferase activity measured following cotransfection with the pBAT8 expression vector alone in the same experiment. The relative *cdx*-3-dependent activation of the proglucagon promoter varied from 5- to 17-fold in different transfections as described previously (16). Each plasmid was transfected in triplicate, and the data shown represent the mean ± S.E. for three different experiments. *, $p < 0.05$ (experimental compared with *cdx*-3 (wild-type) (*wt*)).

terminal deletions, were consistently associated with attenuation of the *cdx*-3-induction of the proglucagon promoter (Fig. 3, *a* and *b*). One of the smallest Cdx-3 amino-terminal mutants containing a deletion within domain A, Δ(8–41), displayed a clearly reduced ability to activate the proglucagon promoter. Progressive deletion of additional *cdx*-3 sequences, as represented by Δ(8–106) and Δ(8–139), produced a further decline in *cdx*-3-dependent transcriptional activation, whereas deletion of the entire *cdx*-3 amino-terminal domain (Δ(8–180)) resulted in complete loss of proglucagon promoter activation (Fig. 3*b*). These observations, taken together with the data demonstrating the functional activity of the first 83 amino acids in the context of a Cdx-3/Pit-1 fusion protein, clearly illustrate the importance of the Cdx-3 amino-terminal region for transcrip-

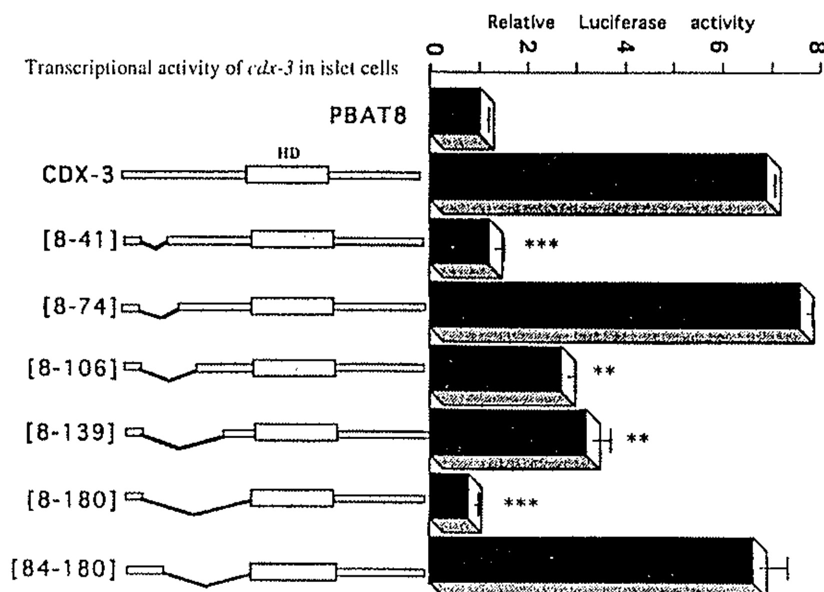

FIG. 4. **Transcriptional properties of *cdx*-3 plasmids in InR1-G9 islet cells.** The *cdx*-3 plasmids (5 μg) were cotransfected into hamster InR1-G9 glucagon-producing islet cells with 5 μg of the −82 proglucagon-luciferase plasmid, and the luciferase activity obtained was normalized to the background luciferase activity measured following cotransfection of the pBAT8 expression vector alone in the same experiment. Each plasmid was transfected in triplicate, and the data shown represent the mean ± S.E. for three different experiments. **, $p < 0.01$; ***, $p < 0.001$ (*cdx*-3 mutants compared with *cdx*-3 (wild-type)).

tional activation.

Carboxyl-terminal regions of transcription factors have also been demonstrated to be functionally important for DNA binding and transcriptional activation (28). Consistent with the data shown in Fig. 1*a*, the Δ(250–313) Cdx-3 molecule lacking the carboxyl-terminal 63 amino acids was associated with reduced levels of transcriptional activation (Fig. 3*c*). In contrast, further deletion of all carboxyl-terminal sequences and the homeodomain resulted in complete loss of *cdx*-3-dependent transcriptional activity (Fig. 3*c*), demonstrating that the Cdx-3 amino-terminal region alone, in the absence of a DNA binding domain, is functionally inactive.

As the proglucagon and *cdx*-2/3 genes are not normally expressed in fibroblasts, we studied the transcriptional properties of *cdx*-3 mutants in islet A cells previously shown to express both the endogenous proglucagon and *cdx*-3 genes (16, 17, 27). Consistent with the data obtained in BHK fibroblasts, the Δ(84–180) mutant did not exhibit a diminution of transcriptional activation in InR1-G9 islet cells (Fig. 4). Similarly, the Δ(8–41) mutant exhibited significantly reduced activity in both BHK and InR1-G9 cells (Figs. 3 and 4), consistent with the functional importance of the proximal amino-terminal region for Cdx-3 activity. In contrast, the Δ(8–74) mutant, encompassing deletions within Cdx-3 domains A and B, exhibited transcriptional activity comparable to wild-type Cdx-3 in islet cells but not in fibroblasts (Fig. 3 *versus* Fig. 4). These observations delineate amino-terminal subdomains exhibiting cell-specific transactivation properties and are consistent with the cell-specific transactivation properties demonstrated for the entire Cdx-2(1–180) amino terminus in fibroblast and intestinal epithelial cell lines (26).

The loss of transcriptional activation following selective deletion of Cdx-3 amino-terminal sequences may be attributable to reduced levels of Cdx-3 expression, incomplete Cdx-3 cytoplasmic to nuclear translocation, diminished binding of Cdx-3 to the proglucagon promoter, and/or impaired activity of one or more transcriptional activation domains. To distinguish between these various mechanisms, we prepared nuclear extracts from BHK and InR1-G9 cells transfected with the *cdx*-3 mutant plasmids. The expression and DNA binding properties of the various Cdx-3 mutant proteins was examined using a combination of Western blotting with Cdx-3 antisera and EMSA analyses (Fig. 5, *a* and *b*) with a proglucagon promoter G1 binding site probe (16, 17). All of the Cdx-3 mutant proteins were expressed and detected in BHK and InR1-G9 nuclear extracts by EMSA and Western blotting (Fig. 5, *a* and *b*, and data not shown). Competition with excess unlabeled Gc probe diminished specific complex formation; however, competition with a heterologous oligonucleotide had no effect (Fig. 5*c*). None of the of the amino-terminal mutants exhibited consistent relative reductions (compared with wild-type Cdx-3 in the levels of protein expression as assessed by Western blotting (data not shown). The lack of available antisera that cross-react with Cdx-3 amino-terminal sequences precluded similar Western blot analyses of the carboxyl-terminal deletions.

To more precisely localize specific amino-terminal amino acid residues that are functionally important for Cdx-3 biological activity, we analyzed the expression and transcriptional activity of a series of single point mutations within the first 83 amino acids of the Cdx-3 amino-terminal region. Despite the conservation of *caudal* amino acid sequences within domain B, mutation of amino acid 28, 69, or 74 had no significant effect on *cdx*-3-dependent proglucagon transcription (Fig. 6). A significant reduction in Cdx-3 transcriptional activity was observed, in both BHK and InR1-G9 cells, with the point mutation closest to the amino terminus of the protein in domain A, a tyrosine-serine substitution at position 14. The reduction in transcriptional activity was not attributable to decreased expression of this mutant protein, as assessed by Western blot analyses (Fig. 6*c*).

Further mutagenesis of sequences within domain A, single point mutations D8A, K9M, and D10A, resulted in Cdx-3 mutants exhibiting significant reductions in proglucagon promoter transcriptional activation in both BHK fibroblasts and InR1-G9 islet cells (Fig. 7*a*). EMSA experiments demonstrated no major impairment of G1 element binding activity (Fig. 7*c*) with the D8A and K9M mutations, although a slight reduction in relative levels of expression of these two mutants (compared

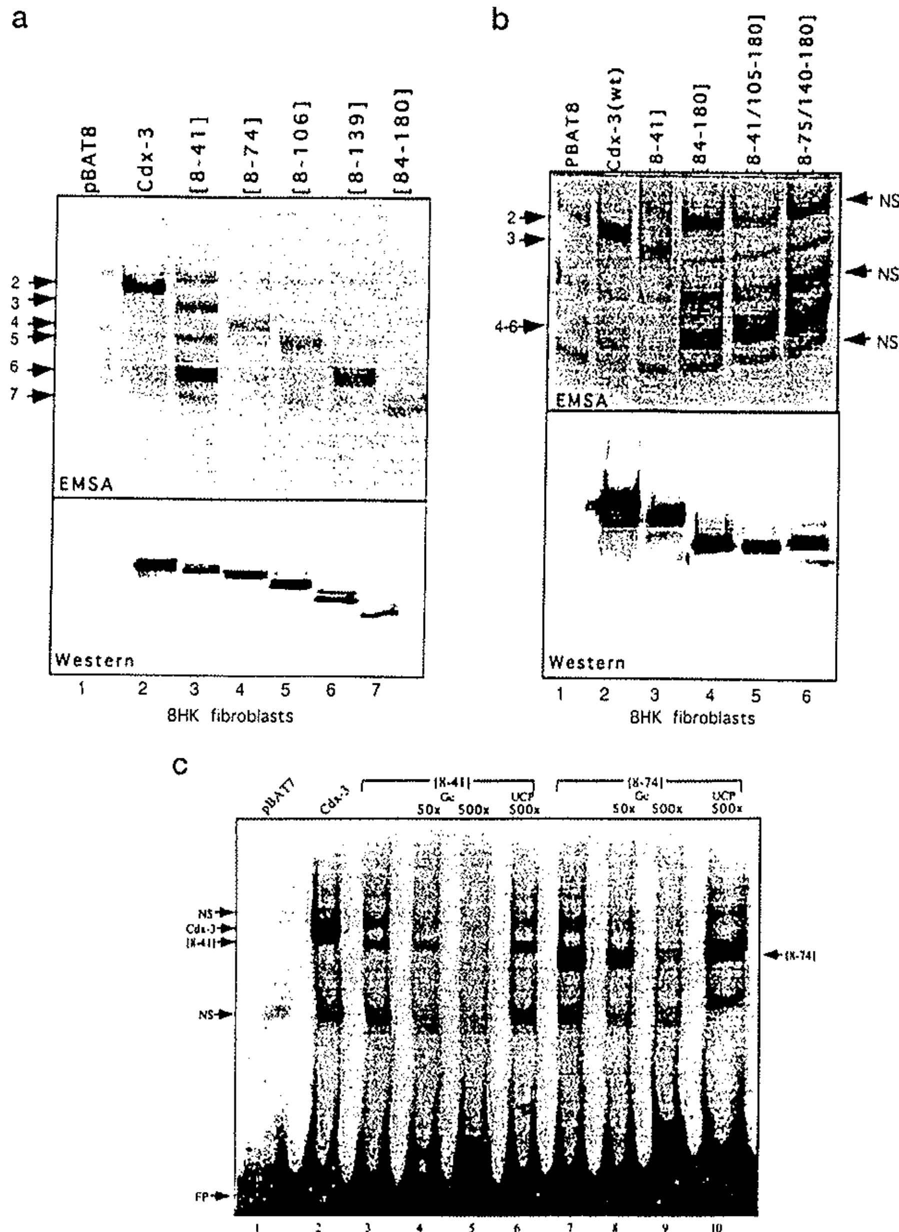

FIG. 5. **Analysis of *cdx*-3 expression by EMSA and Western blotting.** *a* and *b*, *Cdx*-3 plasmids were transfected into BHK cells, and then nuclear extracts were prepared and nuclear protein was utilized for EMSA experiments with a G1 proglucagon promoter probe (16) and for Western blotting with a Cdx-3-specific primary antibody (16). The *numbered arrows* to the *left* of the EMSA panels denote the migration position of the *cdx*-3 mutant-DNA complexes. *NS*, nonspecific complexes detected in the absence of transfected *cdx*-3. *c*, EMSA analysis of Cdx-3(8–41) and Cdx-3(8–74) mutants transfected into BHK cells, using the Gc proglucagon promoter oligonucleotide probe; complex formation was competed for using excess homologous wild-type Gc, or a 27-base pair heterologous oligonucleotide from the uncoupling protein 2 coding region (*UCP*). *NS*, nonspecific complexes; *FP*, free probe. *Arrows* denote the migration positions of the complexes formed with wild-type Cdx-3, Cdx-3(8–41), and Cdx-3(8–74).

with wild-type Cdx-3, was detected by Western blotting (Fig. 7*b*). Similarly, the Cdx-3(D10A) point mutation exhibited a significant reduction in transcriptional activity in both fibroblast and islet cells, without impairment of protein expression or G1 element binding (Fig. 7).

In contrast, the double mutants Cdx-3(D8A/K9M) and Cdx-3(K9M/D10A) and the triple point mutation Cdx-3(D8A/K9M/D10A) resulted in complete elimination of transcriptional activity in islet and fibroblast cells (Fig. 7*a*). Furthermore, Western blotting demonstrated marked reduction in protein expression of the three mutants, reflected in a corresponding decrease in G1 element binding activity (Fig. 7, *b* and *c*). Analysis of the mutationally sensitive Cdx-3 amino-terminal sequences within domain A demonstrated that amino acids 8–10 resembled a diacidic Asp-*X*-Asp motif (Fig. 1*c*) that is remarkably similar to a recently described consensus sequence required for protein export from the endoplasmic reticulum (29).

These observations suggested that the reduction in nuclear expression of Cdx-3 mutant proteins harboring double or triple point mutations in a putative membrane export signal might be attributable to defective Cdx-3 trafficking from the cytoplasm to the nucleus. Analysis of the cellular localization of wild-type *cdx*-3 following transfection of BHK cells demonstrated that the majority of immunoreactive wild-type protein was localized to the nuclear compartment (Fig. 8). In contrast, a marked reduction in the levels of nuclear *cdx*-3(K9M/D10A) was observed in the same experiment, whereas the levels of *cdx*-3 and *cdx*-3(K9M/D10A) in the nuclear membrane fraction were com-

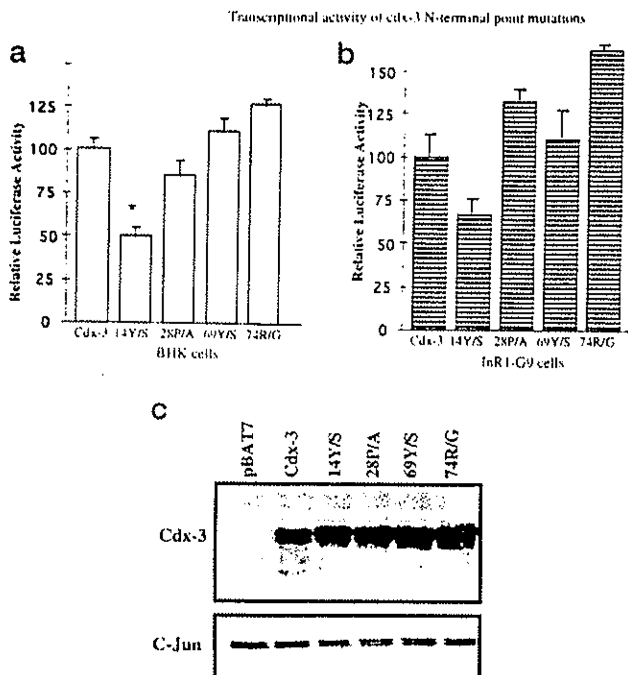

FIG. 6. **Transcriptional properties of Cdx-3 amino-terminal point mutations in BHK and InR1-G9 cells.** *a* and *b*, the transcriptional activity of the individual *cdx*-3 plasmids harboring single point mutations, cotransfected with the −82 proglucagon-luciferase plasmid, was normalized relative to the values obtained following transfection of the wild-type *cdx*-3 plasmid into BHK fibroblasts (*a*) and InR1-G9 islet cells (*b*). *, $p < 0.05$ (point mutations *versus cdx*-3 wild-type). *c*, Western blot analysis of wild-type Cdx-3, Cdx-3 point mutations, and endogenous c-Jun in transfected BHK cells. Due to the expression of endogenous *cdx*-3 in InR1-G9 cells, Western blot analysis of the identical transfected plasmids in InR1-G9 cells could not distinguish between endogenous wild-type Cdx-3 and transfected Cdx-3 point mutations (not shown).

parable. Initial analysis of nonfractionated cell extracts demonstrated markedly reduced levels of cytoplasmic *cdx*-3(K9M/D10A) (data not shown). To localize the cellular distribution of *cdx*-3(K9M/D10A) expression, more detailed cell and nuclear fractionation was performed. Analysis of the cytoplasmic membrane fraction demonstrated low levels of wild-type Cdx-3 but clearly increased levels of the mutant Cdx-3(K9M/D10A) protein (Fig. 8). These findings are consistent with defective cytoplasmic export and arrest of the mutant protein in a cytoplasmic membrane fraction.

DISCUSSION

The *caudal*-related gene designated *cdx*-2 in the mouse and *cdx-3* in the hamster is expressed in both the glucagon-producing cells of the islets and the enteroendocrine cells of the intestine and activates the proglucagon promoter in cell transfection studies (16, 17, 27). The importance of Cdx-2/3 for proglucagon gene transcription was emphasized by experiments demonstrating that Cdx-2/3 activated transcription of the endogenous proglucagon gene in its native chromatin context in transfected islet cells *in vitro* (17). *cdx*-2/3 also regulates an increasing number of genes expressed in the intestine, including the sucrase isomaltase, carbonic anhydrase 1, lactase-phlorizin hydrolase, calbindin-D9K, apolipoprotein B, and intestinal phospholipase A/lysopholipase gene promoters (11, 30–35).

Despite the emerging importance of *caudal* family members for pattern formation and cell differentiation, little is known about the functional relevance of domains outside the homeodomain, such as conserved domains A–C, which are postulated to be potentially important for Caudal protein biological activity (9). Analysis of Cdx-2 structure-function relationships using Cdx-2:GAL4 fusion proteins localized a transcriptional activation domain to the amino-terminal amino acids 15–180 of the Cdx-2 protein (26). Furthermore, the functional activity of the Cdx-2 amino-terminal activation domain was shown to be dependent on the contextual location of promoter binding sites in a cell-specific manner (26). These observations suggested that one or more cellular factors, present in intestinal cells but not fibroblasts, modified the transcriptional activity of Cdx-2 *in vitro*.

The results of our experiments have defined in more detail the specific and minimal amino-terminal Cdx-2/3 sequences capable of functioning as a heterologous activation domain *in vitro*. Our data demonstrate that whereas the amino-terminal 180 amino acids function as an activation domain in a LexA:Cdx fusion protein, as little as the first 83 amino acids of Cdx-3 are capable of functioning as a heterologous transcriptional activation domain when fused to Pit-1 sequences *in vitro*. Furthermore, removal of amino acids 84–180 from the native Cdx-3 protein had no effect on the transcriptional activity of Cdx-3 in BHK fibroblasts or InR1-G9 islet cells. These results provide further evidence that the amino-terminal 83 amino acids in Cdx-3 can function as an activation domain in the context of the native Cdx-3 protein.

Whereas deletions within the first 84 amino acids of Cdx-3 were generally associated with diminution of transcriptional activity in BHK cells, Cdx-3(8–74) did not exhibit a reduction in transcriptional activity in InR1-G9 islet cells. These findings may be potentially explained by the presence of at least two independent modular domains within the first 180 amino acids of Cdx-3 that can functionally compensate for each other in InR1-G9 islet cells, but not in BHK fibroblasts. These observations provide further experimental support for the existence of cell-specific factors (26), differentially expressed in islet *versus* fibroblast cells, that interact with the Cdx-3 activation domain(s) to modify functional activity in a cell-specific manner.

Alignment of the amino-terminal regions of the Caudal homeodomain proteins reveals a number of highly conserved subdomains. designated A–C (Fig. 1c) amino-terminal to the homeobox (9). One of the conserved regions in Caudal proteins, domain A, is represented by the first 14 amino acids (MYVSYLLDKDVSMY) in Cdx-3. Two additional conserved subdomains, designated B (amino acids 68–78) and C (amino acids 161–172) of Cdx-3 (Fig. 1c) are also located within the first 180 amino acids amino-terminal to the homeodomain (9). Chimeric proteins containing either domain B (LexA:Cdx (42–83)) or domain C (LexA:cdx141–180) were not transcriptionally active, suggesting that subdomains B and C alone are not sufficient to confer transcriptional activation in the presence of a heterologous DNA binding domain. Furthermore, deletions in domains B and C, as represented by cdx-3(84–180), did not impair transcriptional activation. In contrast, single point mutations in domain A, as represented by Cdx-3(D10A), exhibited reduced transcriptional activity despite normal levels of nuclear expression and DNA binding, illustrating the importance of domain A for transcriptional activation.

Our results identify an amino-terminal subdomain that is critical for *cdx*-3 expression. Classical nuclear localization signals mediating nuclear import are generally characterized by clusters of basic amino acids, in contrast to the mutationally sensitive diacidic motif identified here for *cdx*-3 (36). Hence, it seems unlikely that sequences within domain A of the Caudal proteins, specifically the DXE motif, constitute a functional nuclear localization signal. Furthermore, domain A is not hydrophobic and would not be predicted to function as a leader

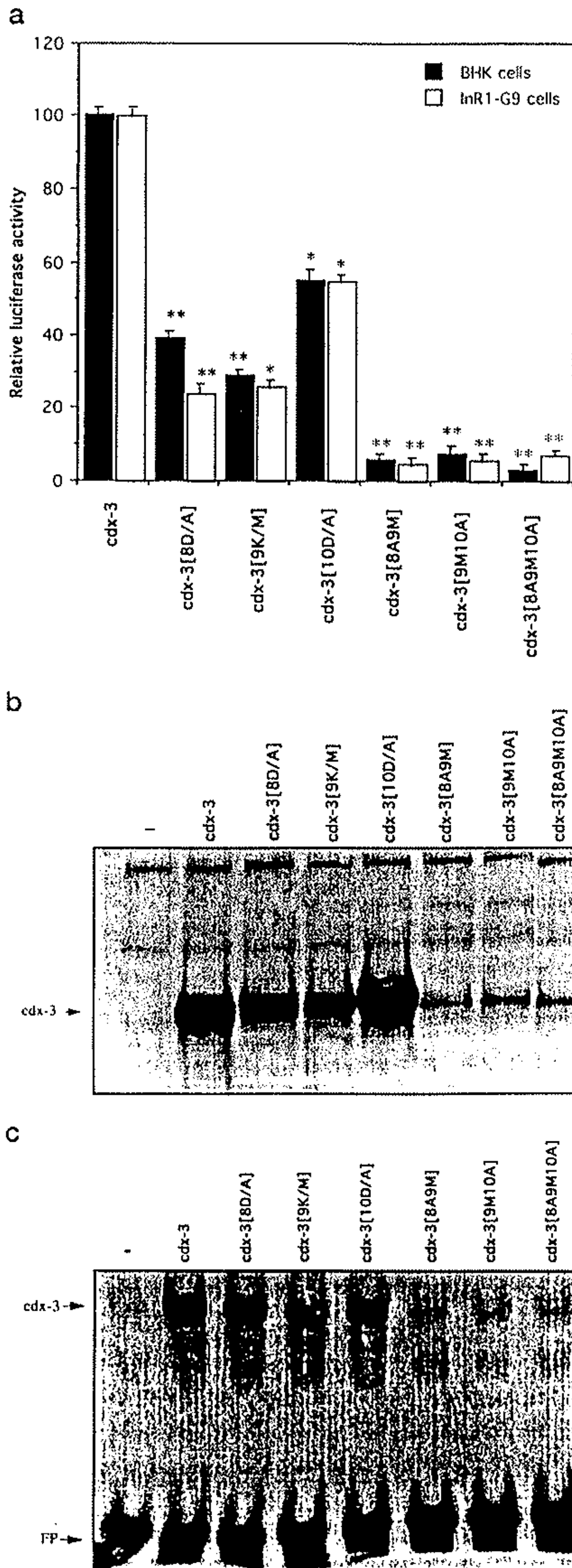

FIG. 7. *a*, relative transcriptional activity of clustered Cdx-3 amino-terminal point mutations in BHK fibroblasts and InR1-G9 islet cells. The transcriptional activity of the *cdx*-3 plasmids harboring single, double, or triple point mutations, cotransfected with the −82 proglucagon-luciferase plasmid, was normalized relative to the values obtained following transfection of the wild-type *cdx*-3 plasmid into BHK fibroblasts or InR1-G9 islet cells. Each plasmid was transfected in triplicate, and the data shown represent the mean ± S.E. of three separate experiments. *, $p < 0.05$; **, $p < 0.05$ (transcriptional activity of point mutations *versus* wild-type *cdx*-3). *b*, Western blot analysis of nuclear extracts following transfection of wild-type *cdx*-3 and *cdx*-3 point mutations in BHK fibroblasts. *c*, EMSA of BHK cell extracts transfected with wild-type and *cdx*-3 amino-terminal point mutations using the proglucagon promoter G1 element probe. − denotes transfection of the expression vector pBAT7 alone. *FP*, free probe.

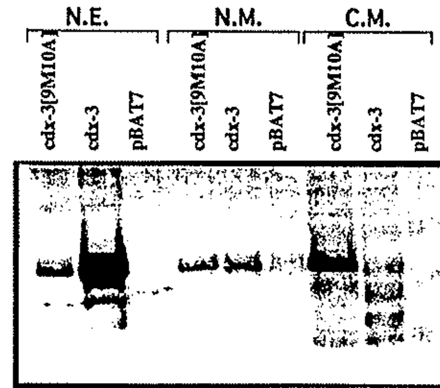

FIG. 8. **Western blot analysis of cytoplasmic and nuclear fractions following transfection of BHK cells with pBAT7 alone or with wild-type *cdx*-3 or *cdx*-3(K9M/D10A).** *N.E.*, nuclear extracts; *N.M.*, nuclear membranes; *C.M.*, cytoplasmic membranes. Cytoplasmic and nuclear membrane fractions were isolated as described under "Materials and Methods."

sequence directing wild-type Cdx-3 to the endoplasmic reticulum compartment (37). Moreover, deletion of the DKD motif from Cdx-3 did not compromise nuclear localization of amino-terminally deleted *cdx*-3 plasmids. The striking homology of the Cdx-3 DXE sequence to a recently described endoplasmic reticulum export signal (29) raises the possibility that Cdx-3 may normally associate with a cytosolic membrane compartment, perhaps in a transient and/or regulated manner. It thus appears that although a functional DXE motif may be necessary for export from a cytoplasmic compartment, it is not absolutely necessary for cytoplasmic to nuclear translocation.

Despite considerable progress in the elucidation of molecular chaperones and mechanisms important for protein folding and export of proteins from the endoplasmic reticulum to the golgi and plasma membranes (37, 38), the molecular machinery that directs proteins such as transcription factors from the cytoplasm to the nucleus remains less well understood (36). The implication from our findings that wild-type Cdx-3 may be transiently associated with a cytoplasmic membrane fraction is not without precedent. For example, both endogenously expressed and transfected engrailed homeodomain proteins have been localized to cytoplasmic caveolaer membrane fractions in subcellular fractionation experiments (39). Furthermore, the sterol regulatory element binding protein-1 transcription factor has also been localized to a cytosolic membrane fraction. Sterol depletion leads to proteolytic generation of a smaller sterol regulatory element binding protein-1 molecule that subsequently translocates from the cytoplasmic membrane compartment to the nucleus (40). Finally, secreted signaling proteins (41) and direct protein-protein interactions (42) may contribute to regulation of cytoplasmic transcription factor sequestration and nuclear import, as shown for the *homothorax* gene that directly interacts with *extradenticle* to induce nuclear translocation; in *homothorax* mutant embryos, *extradenticle* remains transcriptionally inactive due to cytoplasmic sequestration (42).

Our results demonstrate that sequences within *cdx*-3 domain

A are critically important for cytoplasmic export, as mutations in this region lead to arrest in a cytoplasmic membrane fraction and subsequent compromised nuclear localization. Given the increasing evidence that even partial reductions in *cdx-2/3* expression are associated with changes in cell physiology (12), proliferation (14), and tissue development (14), mutations in the Cdx-2/3 amino-terminal region that affect protein compartmentalization or transcriptional activation may be expected to produce a significant biological phenotype *in vivo*.

REFERENCES

1. Mlodzik, M., Fjose, A., and Gehring, W. A. (1985) *EMBO J.* **4**, 2961–2969
2. Macdonald, P. M., and Struhl, G. (1986) *Nature* **324**, 537–545
3. Mlodzik, M., Gibson, G., and Gehring, W. J. (1990) *Development* **109**, 271–277
4. James, R., and Kazenwadel, J. (1991) *J. Biol. Chem.* **266**, 3246–3251
5. Hu, Y., Kazenwadel, J., and James, R. (1993) *J. Biol. Chem.* **268**, 27214–27225
6. Bonner, C. A., Loftus, S. K., and Wasmuth, J. J. (1995) *Genomics* **28**, 206–211
7. Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J., and Gruss, P. (1988) *Genes Dev.* **2**, 1647–1654
8. Horn, J. M., and Ashworth, A. (1995) *Hum. Mol. Genet.* **4**, 1041–1047
9. Gamer, L. W., and Wright, C. V. E. (1993) *Mech. Dev.* **43**, 71–81
10. Doll, U., and Niessing, J. (1996) *Eur. J. Cell Biol.* **70**, 260–268
11. Suh, E., and Traber, P. G. (1996) *Mol. Cell. Biol.* **16**, 619–625
12. Lorentz, O., Duluc, I., De Arcangelis, A., Simon-Assmann, P., Kedinger, M., and Freund, J.-N. (1997) *J. Cell Biol.* **139**, 1553–1565
13. Subramanian, V., Meyer, B. I., and Gruss, P. (1995) *Cell* **83**, 641–653
14. Chawengsaksophak, K., James, R., Hammond, V. E., Kontgen, F., and Beck, F. (1997) *Nature* **386**, 84–87
15. Joly, J. S., Maury, M., Joly, C., Duprey, P., Boulekbache, H., and Condamine, H. (1992) *Differentiation* **50**, 75–87
16. Jin, T., and Drucker, D. J. (1996) *Mol. Cell. Biol.* **16**, 19–28
17. Jin, T., Trinh, D. K. Y., Wang, F., and Drucker, D. J. (1997) *Mol. Endocrinol.* **11**, 203–209
18. German, M. S., Wang, J., Chadwick, R. B., and Rutter, W. J. (1992) *Genes Dev.* **6**, 2165–2176
19. Lew, A. M., and Elsholtz, H. P. (1995) *J. Biol. Chem.* **270**, 7156–7160
20. Elsholtz, H. P., Lew, A. M., Albert, P. R., and Sundmark, V. C. (1991) *J. Biol. Chem.* **266**, 22919–22925
21. Chen, X., Liu, W., Quinto, I., and Scala, G. (1997) *Nucleic Acids Res.* **25**, 682–684
22. Drucker, D. J., Philippe, J., and Mojsov, S. (1988) *Endocrinology* **123**, 1861–1867
23. Wang, M., and Drucker, D. J. (1995) *J. Biol. Chem.* **270**, 12646–12652
24. Marmiroli, S., Ognibene, A., Bavelloni, A., Cinti, C., Cocco, L., and Maraldi, N. M. (1994) *J. Biol. Chem.* **269**, 13–16
25. Kaufmann, S. H., Gibson, W., and Shaper, J. H. (1983) *J. Biol. Chem.* **258**, 2710–2719
26. Taylor, J. K., Levy, T., Suh, E. R., and Traber, P. G. (1997) *Nucleic Acids Res.* **25**, 2293–2300
27. Laser, B., Meda, P., Constant, I., and Philippe, J. (1996) *J. Biol. Chem.* **271**, 28984–28994
28. Stark, M. R., and Johnson, A. D. (1994) *Nature* **371**, 429–432
29. Nishimura, N., and Balch, W. E. (1997) *Science* **277**, 556–558
30. Suh, E., Chen, L., Taylor, J., and Traber, P. G. (1994) *Mol. Cell. Biol.* **14**, 7340–7351
31. Drummond, F., Sowden, J., Morrison, K., and Edwards, Y. H. (1996) *Eur. J. Biochem.* **236**, 670–681
32. Troelsen, J. T., Mitchelmore, C., Spodsberg, N., Jensen, A. M., Noren, O., and Sjostrom, H. (1997) *Biochem. J.* **322**, 833–838
33. Lee, S.-Y., Nagy, B. P., Brooks, A. R., Wang, D.-M., Paulweber, B., and Levy-Wilson, B. (1996) *J. Biol. Chem.* **271**, 707–718
34. Lambert, M., Colnot, S., Suh, E., L'Horset, F., Blin, C., Calliot, M. E., Raymondjean, M., Thomasset, M., Traber, P. G., and Perret, C. (1996) *Eur. J. Biochem.* **236**, 778–788
35. Taylor, J. K., Boll, W., Levy, T., Suh, E., Siang, S., Mantei, N., and Traber, P. G. (1997) *DNA Cell Biol.* **16**, 1419–1428
36. Gorlich, D., and Mattaj, I. W. (1996) *Science* **271**, 1513–1518
37. Schatz, G., and Dobberstein, B. (1996) *Science* **271**, 1519–1526
38. Kim, P. S., and Arvan, P. (1998) *Endocr. Rev* **19**, 173–202
39. Joliot, A., Trembleau, A., Raposo, G., Calvet, S., Volovitch, M., and Prochiantz, A. (1997) *Development* **124**, 1865–1875
40. Wang, X., Sato, R., Brown, M. S., Hua, X., and Goldstein, J. L. (1994) *Cell* **77**, 53–62
41. Mann, R. S., and Abu-Shaar, M. (1996) *Nature* **383**, 630–633
42. Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M., and Mann, R. S. (1997) *Cell* **91**, 171–183