Characterization of Somatostatin Transactivating Factor-1, a Novel Homeobox Factor That Stimulates Somatostatin Expression in Pancreatic Islet Cells

James Leonard, Bernard Peers, Ted Johnson, Kevin Ferreri, Soon Lee, and Marc R. Montminy*

The Clayton Foundation Laboratories for Peptide Biology
The Salk Institute for Biological Studies
La Jolla, California 92037

The endocrine pancreas consists of several differentiated cell types that are distinguished by their selective expression of peptide hormones such as insulin, glucagon, and somatostatin. Although a number of homeobox-type factors have been proposed as key regulators of individual peptide genes in the pancreas, their cellular distribution and relative abundance remain uncharacterized. Also, their overlapping DNA binding specificities have further obscured the regulatory functions these factors perform during development. In this report we characterize a novel homeobox-type somatostatin transactivating factor termed STF-1, which is uniformly expressed in cells of the endocrine pancreas and small intestine. The 283-amino acid STF-1 protein binds to tissue-specific elements within the somatostatin promoter and stimulates somatostatin gene expression both *in vivo* and *in vitro*. Remarkably, STF-1 comprises the predominant tissue-specific element-binding activity in nuclear extracts from somatostatin-producing pancreatic islet cells, suggesting that this protein may have a primary role in regulating peptide hormone expression and specifying endocrine cell lineage in the developing gut. (Molecular Endocrinology 7: 1275–1283, 1993)

INTRODUCTION

The endocrine pancreas consists primarily of three cell types, which are distinguished by their selective expression of glucagon, insulin, or somatostatin. Embryologically derived from the small intestine (1), these so-called pancreatic islet cells may retain regulatory pathways that originate in the gut to direct expression of the same peptide hormone genes. In its early stages, the fetal endocrine pancreas is populated by pluripotential cells that can coexpress somatostatin, insulin, and glucagon (2). As these stem cells mature, their endocrine repertoire becomes restricted to a single peptide gene, suggesting that whereas common factors may initially regulate all three genes, distinct nuclear proteins must subsequently specify peptide production in individual mature islet cell types.

We (3) and others (4) have previously observed that somatostatin gene expression in the pancreatic islet cell line Tu-6 requires a tissue-specific promoter element (TSE) that operates in concert with the cAMP response element (CRE) to provide high level constitutive activity. TSE-like sequences are reiterated three times over a 500-basepair region of the somatostatin promoter, with the promoter proximal TSEs, located at −300 and −100, being the most active. The somatostatin TSEs contain a canonical TAAT motif that is generally recognized by homeobox-type proteins. The homeobox factor ISL-1, for example, can bind to the TSE and regulate somatostatin expression in the rat insulinoma cell line RIN 5AH. However, ISL-1 appears to comprise only a negligible fraction of TSE-binding activity in extracts of somatostatin-producing Tu-6 cells (our unpublished observations) (5), prompting us to examine whether other homeobox proteins might be more directly involved in regulating somatostatin expression in these cells.

In this report, we describe a novel somatostatin transactivating factor termed STF-1, whose expression is restricted to endocrine cells within the small intestine and pancreatic islets, sites of abundant somatostatin production in adult tissues. In pancreatic Tu-6 cells, STF-1 protein accounts for nearly all of the high affinity TSE-binding activity, and STF-1 stimulates somatostatin expression both *in vitro* and *in vivo*, demonstrating that this factor may act as a transcription factor. As STF-1 expression is not confined to somatostatin-producing D-cells of the gut and pancreas, but is distributed throughout the pancreatic islets and small intestine, this

0888-8809/93/1275-1283$03.00/0
Molecular Endocrinology

factor may be more generally involved in promoting peptide hormone expression at these sites.

RESULTS

Isolation of STF-1 cDNA Clones

To isolate homeobox factors that might regulate somatostatin gene expression in Tu-6 cells, we prepared degenerate sense and antisense oligos corresponding to the most highly conserved regions of the homeodomain, amino acids 16–20 (LEKEF) and amino acids 47–51 (IWFQN) of the homeobox, respectively (6, 7). Using these primers for polymerase chain reaction (PCR) amplification of sequences in a Tu-6 cDNA library, we obtained a DNA fragment of the predicted size that was subcloned into Bluescript SK II. Of six recombinant clones analyzed, five corresponded to the same homeodomain fragment, suggesting that this cDNA might encode the most prevalent homeobox type protein in Tu-6 cells. Using the PCR fragment as hybridization probe, we obtained a full-length 1.6-kb cDNA clone from a Tu-6 cDNA library. This cDNA encoded a novel protein of 283 amino acids that we have termed STF-1 (Fig. 1A).

Within the homeobox region (amino acids 146–205), STF-1 shares considerable homology with the *Antennapedia* family of proteins (Fig. 1B). The preferred recognition sequence for *Antennapedia* is CTAATG (7), a motif that is also contained in the promoter proximal somatostatin TSE-I. The STF-1 homeobox is clearly most homologous to the leech factor Htr-A2 (88%) (8), a protein of unknown function, and to *Xenopus* XlHbox 8 (100% of the available sequence) (9), a protein expressed exclusively in small intestine and exocrine pancreas. These three proteins share several features that distinguish them from other homeobox proteins and that may impart unique DNA-binding characteristics to these factors. Most notably, each protein contains a histidine at position 44 within the recognition helix that replaces an otherwise conserved glutamine. Consequently, we suggest that these factors may be grouped into a family of proteins referred to as SHX to indicate the first letter for each protein.

Expression of STF-1 RNA

To determine whether STF-1 production is limited to endocrine cell types associated with somatostatin production, we performed RNAse protection studies on RNAs obtained from a variety of cell lines and tissues (Fig. 2). Using a labeled STF-1 antisense RNA probe, we observed STF-1 RNA in both Tu-6 and RIN5 AH pancreatic islet cells, but no detectable STF-1 RNA in nonislet cell lines, including PC12, JEG-3, COS, HT 22, HeLa (not shown), and others. Of the 11 rat tissues we examined, only pancreas and small intestine contained STF-1 RNA, demonstrating that the corresponding STF-1 protein may be highly restricted to endocrine cells of the gut and pancreas.

As the pancreas contains both exocrine and endocrine cellular components, we sought to determine the site of STF-1 RNA production by *in situ* hybridization (Fig. 3A). Using an ^{35}S-labeled STF-1 antisense riboprobe, we could detect STF-1 RNA in islets, but not in surrounding exocrine acinar cells. Within the islet, the hybridization signal was evenly distributed over all cells. Because only 10–20% of islet cells produce somatostatin, other cell types, such as insulin-containing β-cells and glucagon-producing α-cells, may also express this factor. STF-1 RNA was present in most epithelial cells of the small intestine (Fig. 3B), but only a small number of these was found to produce somatostatin. Thus, in both tissues, somatostatin-producing cells appear to account for only a small subset of those expressing STF-1.

STF-1 Binds Functional Elements in the Somatostatin Promoter

Within the somatostatin promoter, two related tissue-specific elements, termed TSE-I and -II, promote somatostatin expression in pancreatic islet Tu-6 cells (3) (our unpublished data). To determine whether STF-1 could bind to these functionally defined elements, we prepared recombinant STF-1 protein from *E. coli*-transformed with a prokaryotic GST-STF-1 expression plasmid. After purification on glutathione-agarose beads, the GST-STF-1 fusion protein was cleaved with thrombin, permitting retrieval of a 161-amino acid STF-1 polypeptide fragment extending from amino acids 124–283. Using two somatostatin promoter fragments as probes in DNAse-I protection assays, we observed discreet footprints over both TSE-I and TSE-II sites (Fig. 4). Protection at both sites coincided with the footprinting patterns obtained using crude Tu-6 nuclear extracts (Fig. 4) (3). The TSE-II site was completely protected at far lower concentrations of STF-1, suggesting that this site might bind this protein with higher affinity than the TSE-I site.

We also evaluated the DNA-binding properties of STF-1 with gel mobility shift assays using labeled TSE-I and TSE-II oligos (Fig. 5). Binding of recombinant STF-1 protein to either site was readily displaced by an excess of unlabeled wild-type TSE-I or TSE-II competitor DNAs. However, mutant versions of TSE-I and TSE-II with substitutions in the TAAT recognition motif could not compete for STF-1 binding. As with footprinting studies, the affinity of STF-1 protein for TSE-II appeared to be higher than that for TSE-I.

To determine the abundance of STF-1 relative to other DNA-binding activities in somatostatin-expressing cells, we performed gel shift experiments with Tu-6 nuclear extracts (Fig. 6). Using the high affinity TSE-II probe, we observed three complexes, termed C1, C2, and C3, based on their relative mobility. Complexes C1 and C3 appeared with low amounts of extract, whereas complex C2 emerged when higher concentrations of

A

```
  1 gaattccacgcggctggtggtgataggagccatgttttctgcgtgctctgtccgaggtgctgaaagaactccaggcagattcacctggaaggaccctgaa 100

101 acaaggcttccaggggaaacacgggggatccggggaccggcagcggcagcgggaggggctggaggaaggtccgcgctctctatcagcaatgtgccaccct 200

201 gcccagagcagtggagaactgtcaaagcgatctggggtggcgctgagagtccgtgagctgcccagcgccttaaggcctggcttgtagctccctaccccgg 300

301 gctgccggccccgaagtgccggctgccacc ATG AAT AGT GAG GAG CAG TAC TAC GCG GCC ACA CAG CTC TAC AAG GAC CCG   381
  1                             M   N   S   E   E   Q   Y   Y   A   A   T   Q   L   Y   K   D   P    17

382 TGC GCA TTC CAG AGG GGT CCG GTG CCA GAG TTC AGT GCT AAT CCC CCT GCG TGC CTG TAC ATG GGC CGC CAG CCC  456
 18 C   A   F   Q   R   G   P   V   P   E   F   S   A   N   P   P   A   C   L   Y   M   G   R   Q   P    42

457 CCA CCT CCG CCG CCA CCC CAG TTT GCA GGC TCG CTG GGA ACG CTG GAA CAG GGA AGT CCC CCG GAC ATC TCC CCA  531
 43 P   P   P   P   P   P   Q   F   A   G   S   L   G   T   L   E   Q   G   S   P   P   D   I   S   P    67

532 TAC GAA GTG CCC CCG CTC GCC GAT GAC CCG GCT GGC GCG CAC CTC CAC CAC CAC CTC CCA GCT CAG CTC GGG CTC  606
 68 Y   E   V   P   P   L   A   D   D   P   A   G   A   H   L   H   H   H   L   P   A   Q   L   G   L    92

607 GCC CAT CCA CCT CCC GGA CCT TTC CCG AAT GGA ACC GAG ACT GGG GGC CTG GAA GAG CCC AGC CGC GTT CAT CTC  681
 93 A   H   P   P   P   G   P   F   P   N   G   T   E   T   G   G   L   E   E   P   S   R   V   H   L   117

682 CCT TTC CCG TGG ATG AAA TCC ACC AAA GCT CAC GCG TGG AAA AGC CAG TGG GCA GGA GGT GCA TAC GCA GCA GAA  756
118 P   F   P   W   M   K   S   T   K   A   H   A   W   K   S   Q   W   A   G   G   A   Y   A   A   E   142

757 CCG GAG GAG AAT AAG AGG ACC CGT ACA GCC TAC ACT CGG GCC CAG CTG CTG GAG CTG GAG AAG GAA TTC TTA TTT  831
143 P   E   E   N   K   R   T   R   T   A   Y   T   R   A   Q   L   L   E   L   E   K   E   F   L   F   167

832 AAC AAA TAC ATC TCC CGG CCT CGC CGG GTG GAG CTG GCA GTG ATG CTC AAC TTG ACT GAG AGA CAC ATC AAA ATC  906
168 N   K   Y   I   S   R   P   R   R   V   E   L   A   V   M   L   N   L   T   E   R   H   I   K   I   192

907 TGG TTC CAA AAC CGT CGC ATG AAG TGG AAG AAA GAG GAA GAT AAG AAA CGT AGT AGC GGG ACA ACG AGC GGG GGC  981
193 W   F   Q   N   R   R   M   K   W   K   K   E   E   D   K   K   R   S   S   G   T   T   S   G   G   217

982 GGT GGG GGC GAA GAG CCG GAG CAG GAT TGT GCC GTA ACC TCG GGC GAG GAG CTG CTG GCA TTG CCA CCG CCA CCA 1056
218 G   G   G   E   E   P   E   Q   D   C   A   V   T   S   G   E   E   L   L   A   L   P   P   P   P   242

1057 CCT CCC GGA GGT GCT GTG CCC TCA GGC GTC CCT GCT GCT GCC CGG GAG GGC CGA CTG CCT TCC GGC CTT AGT GCG 1131
 243 P   P   G   G   A   V   P   S   G   V   P   A   A   A   R   E   G   R   L   P   S   G   L   S   A   267

1132 TCC CCA CAG CCC TCC AGC ATC GCG CCA CTG CGA CCG CAG GAA CCC CGG TGA ggaccgcaggctgagggtgagcgggtctggga 1214
 268 S   P   Q   P   S   S   I   A   P   L   R   P   Q   E   P   R   *                                  284

1215 cccagagtgcggacatgggcatgggcccgggcagctggataagggaggggatcatgaggcttaacctaaacgccacacaaggagaacattcttcttgggg 1314

1315 gcacaagagccagttgggtatagccagcgagatgctggcagacctctgggaaaaaaaaagacccgagcttctgaaaactttgaggctgcctctcgtgcca 1414

1415 tgtgaaccgccaggtctgcctctgggactctttcctgggaccaatttagagaatcaggctcccaactgaggacaatgaaaaggttacaaacttgagcggt 1514

 515 cccataacagccaccaggcgagctggaccgggtgcctttgactggtcggccgagcaatctaaggttgagaataaagggagctgtttgaggtt         1606
```

B

ALIGNMENT OF SHX-TYPE HOMEODOMAINS

	1	44	60	Stf-1 Homology
Stf-1	NKRT RTAYTRAQLL ELEKEFLFNK YISRPRRVEL AVMLNLTERHIKIWFQNRRM KWKKEEDKKR			
Xhox8	EPLFNK YISRPRRVEL AVMLNLTERHIKIWFQNRRM KWKKEEDKKR			42/42=100%
Htr-A2	NKRT RTAYSRSQLL ELEKEPHFDK YISRPRRVEL ASSLNLTERHIKIWFQNRRM KWKKME			53/60= 88%
Antp	RKRG RQTYTRYQTL ELEKEFHFNR YLTRRRRIEI AHALCLTERQIKIWFQNRRM KWKKEN			42/60= 70%
Dfd	PKRQ RTAYTRHQIL ELEKEFHYNR YLTRRRRIEI AHTLVLSERQIKIWFQNRRM KWKKDN			41/60= 68%
Lab	NNSG RTNFTNKQLT ELEKEPHFNR YLTRARRIEI ANTLQLNETQVKIWFQNRRM KQKKRV			35/60= 58%
AbdB	VRKK RKPYSKFQTL ELEKEFLFNA YVSKQKRWEL ARNLQLTERQVKIWFQNRRM KNKKNS			36/60= 60%
En	EKRP RTAFSSEQLA RLKREFNENR YLTERRRQQL SSELGLNEAQIKIWFQNKRA KIKKST			29/60= 48%
Hox2.4	RRRG RQTYSRYQTL ELEKEFLFNP YLTRKRRIEV SHALGLTERQVKIWFQNRRM KWKKEN			39/60= 65%
Cad	KDKY RVVYTDFQRL ELEKEYCTSR YITIRRKSEL AQTLSLSERQVKIWFQNRRA KERTSN			28/60= 47%

Fig. 1. Characterization of STF-1 cDNA

A, The nucleotide and corresponding amino acid sequence of STF-1 cDNA obtained from a Tu-6 cDNA library. Numbers on each side indicate nucleotide (*top*) or amino acid (*bottom*) position. 5′- and 3′-flanking sequences are shown in *lowercase*, and the coding sequence is *capitalized*. The corresponding amino acid sequence is shown in single letter code, with the stop codon indicated by an *asterisk*. The 60-amino acid homeodomain is *underlined*. B, Comparison of STF-1 homeodomain with other homeobox proteins. Single letter amino acid sequence starts at position 1 of the homeobox shown. Areas of shared homology with STF-1 are *shaded*. The percent homology with STF-1 homeobox is indicated.

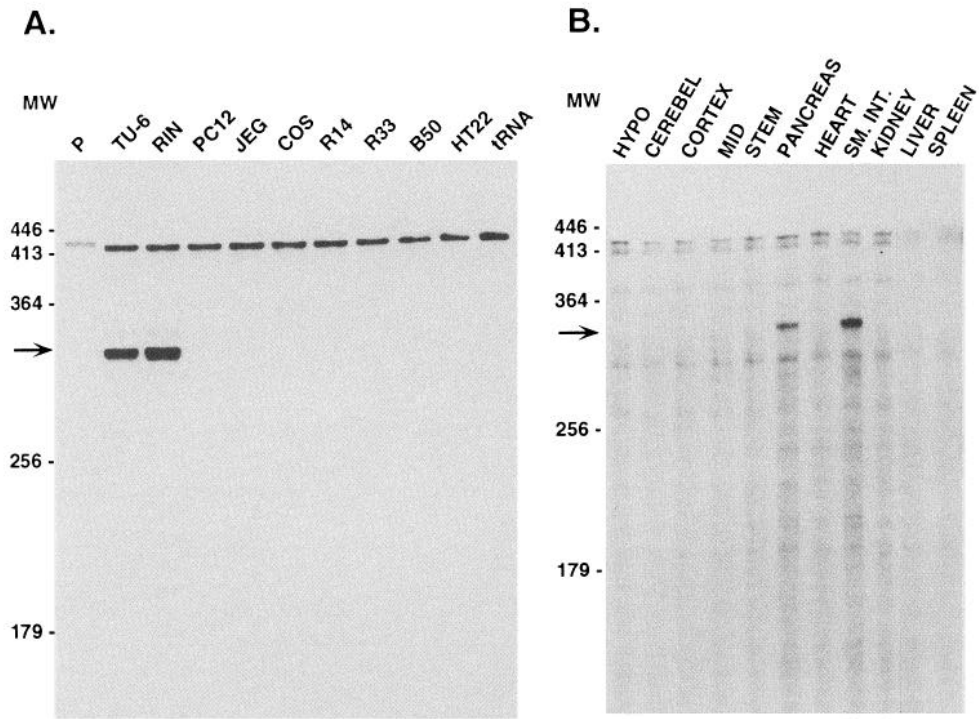

Fig. 2. Expression of STF-1 RNA Is Restricted to Pancreatic Islets and Small Intestine

A, RNAse protection analysis using STF-1 riboprobe on total RNA (30 μg) recovered from cell lines, indicated *above each lane*. P, Probe alone; Tu-6 and RIN, rat insulinoma cell lines; PC12, rat pheochromocytoma cells; JEG, placental choriocarcinoma cells; R13, R33, B50, and HT22, rat retinal cell lines; tRNA, control reaction supplemented with transfer RNA. The *arrow* points to the expected position of the 318-nucleotide RNAse-protected fragment. The 420-base fragment corresponds to undigested probe. B, RNAse protection analysis of various rat tissue RNAs, as indicated *over each lane*. Hypo, Hypothalamus; Cerebel, cerebellum; Mid, midbrain; Stem, brainstem; Sm. Int., small intestine.

extract were employed. C1 and C2 would appear to represent high affinity complexes, as both were displaced by wild-type, but not mutant, TSE-I and TSE-II oligos. As neither C1 nor C2 was observed in nonislet cell extracts such as Hela (not shown), these complexes would also appear to be tissue specific.

To determine whether these complexes might contain factors related to STF-1, we developed a rabbit polyclonal antiserum against a synthetic STF-1 peptide extending from amino acids 196–214 (Figs. 5 and 6). This antiserum could specifically recognize the recombinant STF-1 protein by Western blot analysis (Fig. 5B). When tested in gel mobility shift assays, the STF-1 antiserum specifically inhibited the binding of recombinant STF-1, but not ISL-1 protein (not shown), to the TSE probes (Fig. 5, lane 8).

In Tu-6 extracts, STF-1 antiserum specifically recognized a 49-kilodalton (kDa) protein in nuclear and cytoplasmic extracts (Fig. 6B). Although the mol wt of this band is quite different from the predicted mass of STF-1 (31 kDa), the 49-kDa immunoreactive product comigrates with the *in vitro* translation product from reticulocyte lysates programmed with STF-1 RNA (Fig. 6B). As predicted from Western blot data showing STF-1 protein in both nuclear and cytoplasmic fractions, complex C1 was observed in both extracts (cytoplasmic and nuclear), suggesting that this protein may leak out during extract preparation. When preincubated with Tu-6 nuclear extract (Fig. 6), the STF-1 antiserum completely abolished C1 and C2 complexes, but had no effect on C3. As preimmune serum had no effect on any of these complexes, our results suggest that STF-1 protein accounts for the majority of TSE binding activity in Tu-6 cells. Moreover, complex C1 had the same relative mobility as the full-length recombinant STF-1 protein-TSE complex, further suggesting that C1 contains STF-1 (not shown). The appearance of C2 with increasing concentrations of Tu-6 nuclear extract indicates that this complex could be due to a dimeric form of STF-1. Indeed the TSE-II site contains two TAAT motifs, which may encourage cooperative binding between two homeodomain monomers at this site.

STF-1 Activates Transcription from the Somatostatin Promoter

To determine whether STF-1 could directly stimulate somatostatin transcription, we performed *in vitro* transcription assays using recombinant STF-1 protein (Fig. 7A). Previous studies showing that high level expression of somatostatin in Tu-6 cells also requires CRE-binding protein (CREB) activity (3) prompted us to eval-

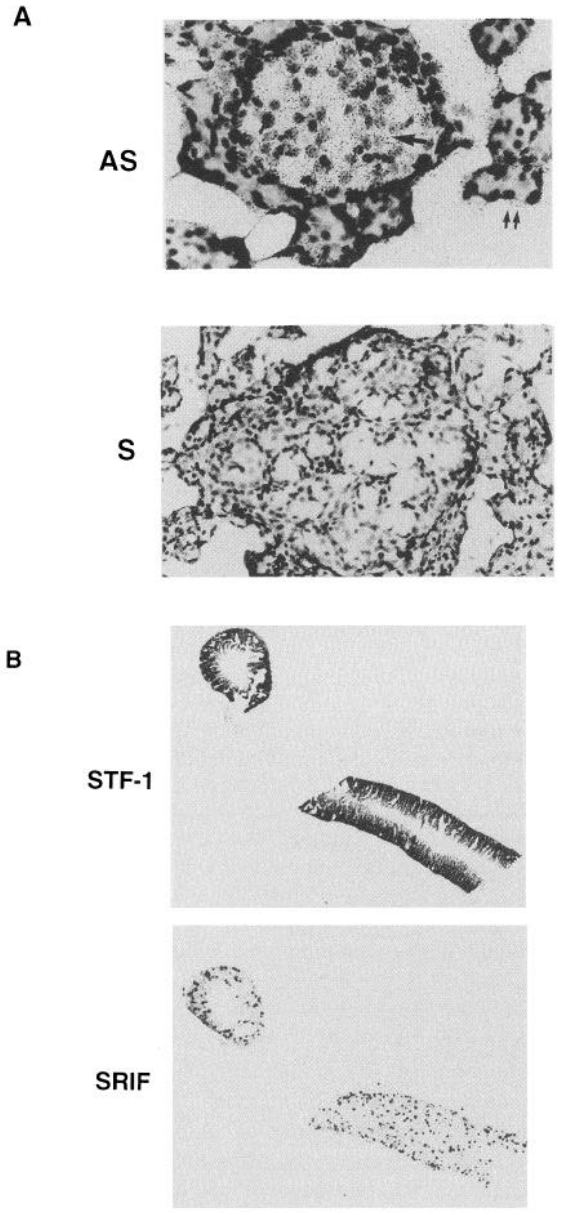

Fig. 3. Expression of STF-1 RNA in Pancreas and Small Intestine

A, *In situ* hybridization analysis of 30-μm sections from rat pancreas. A cross-section of pancreatic islet with surrounding acinar cells is shown in both micrographs. AS (*top*) and S (*bottom*) indicate hybridization with antisense or sense STF-1 RNA probes, respectively. *Small double arrows* point to exocrine acinar cells in the *top micrograph*. The *large arrow* in same micrograph points to hybridization signal limited to pancreatic islet cells. Magnification, ×1800. B, Comparison of STF-1 (*top*) and somatostatin (SRIF; *bottom*) RNA expression in small intestine. Film autoradiogram shows cross-section and longitudinal section of small intestine for both.

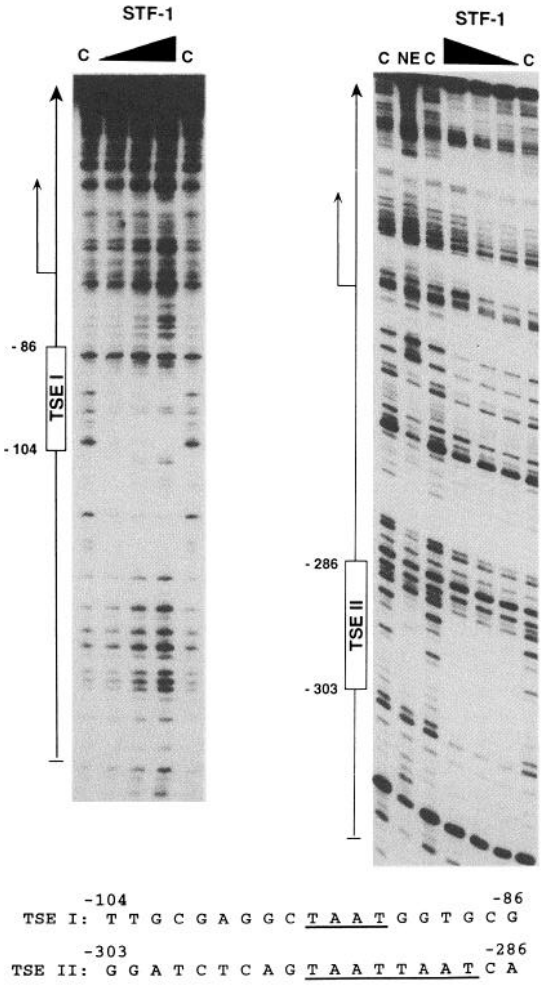

Fig. 4. DNase-I Footprinting Analysis of STF-1-Binding Activity on the Somatostatin Promoter

Left, Autoradiogram of DNAse-I protection assay using somatostatin promoter fragment extending from −141 to 55. The *stick diagram* shows functionally defined TSE-I element with nucleotide position relative to the transcriptional start site, as indicated. *Arrows* show the direction of transcription. C, Control reaction with no extract. The *crescendo sign* indicates increasing amounts of recombinant STF-1 (amino acids 124–283). *Right*, Footprinting assay using 280-nucleotide somatostatin promoter fragment extending from −350 to −71. NE, Tu-6 nuclear extract. The position of TSE-II element on somatostatin promoter with nucleotide boundaries is indicated on the *stick diagram*. *Bottom*, Nucleotide sequence of the STF-1 protected sequences, termed TSE-I and TSE-II elements, in rat somatostatin gene, with conserved TAAT motifs involved in homeodomain recognition *underlined*.

uate STF-1 and CREB proteins separately and in unison. HeLa nuclear extracts lack detectable amounts of STF-1 protein, permitting us to test this factor without interference from the endogenous protein. Using a somatostatin promoter template containing four TSE-I sites (4 × TSE) inserted up-stream of the somatostatin CRE, we observed marked induction of somatostatin transcription after the addition of purified STF-1 protein. By contrast, STF-1 had no stimulatory effect on the α-globin control promoter or on a somatostatin template lacking the TSE site (−TSE) at any level of protein tested. Nevertheless, purified CREB protein stimulated transcription from both ×TSE and 4 × TSE-I templates.

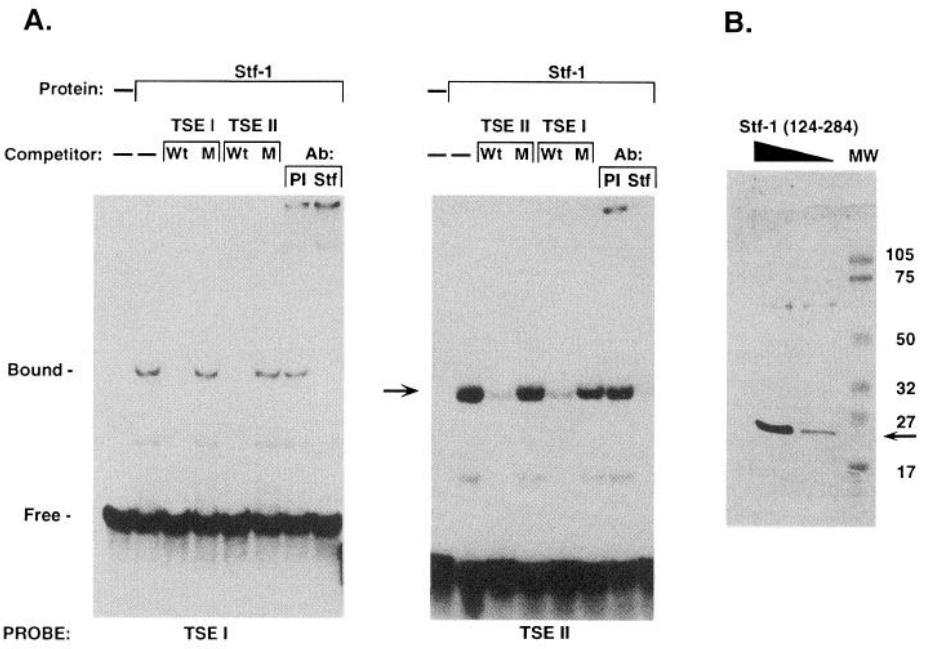

Fig. 5. STF-1 Binds to TSE-I and TSE-II Sequences with High Affinity

A, STF-1-binding activity using somatostatin TSE-I (*left*) or TSE-II (*right*) probes. Protein indicates the presence (Stf-1) or absence (−) of recombinant STF-1 protein fragment (amino acids 124–284) in appropriate lanes. Competitor, Absence (−) or presence (TSE-I and TSE-II) of unlabeled double stranded competitor DNAs (20 ng) in gel shift reactions. Wild-type (WT) or mutant (M) competitor oligos are indicated *over each lane*. Ab, Preimmune (PI) or STF-1 (Stf) antibody added to gel shift reactions. Bound and free indicate the positions of protein-DNA complexes and free DNA probe, respectively. The *arrow* indicates the position of the protein-DNA complex with TSE-II probe. B, Western blot analysis of recombinant STF-1 protein fragment (amino acids 124–283). The *decrescendo sign* indicates a decreasing amount of STF-1 protein. MW, Prestained mol wt marker, with relative mass (in kilodaltons) indicated alongside.

To determine whether STF-1 could also stimulate somatostatin transcription *in vivo*, we inserted the STF-1 cDNA into a cytomegalovirus (CMV) expression vector. When examined in HeLa cells, the CMV-STF-1 expression plasmid stimulated both TSE-I and TSE-II somatostatin reporter plasmids about 12-fold, whereas a parent somatostatin promoter plasmid lacking these sites was only modestly affected (Fig. 7B). CMV-STF-1 also stimulated TSE-I and TSE-II reporter activity upon cotransfection into PC12 cells (not shown), but the −TSE plasmid showed no induction. These results indicate that STF-1 can specifically stimulate transcription from the somatostatin promoter in a cell-type independent manner.

DISCUSSION

Our results suggest a significant role for STF-1 as a tissue-specific regulator of somatostatin transcription. STF-1 expression is highly restricted to endocrine cell types within the pancreas and small intestine, where it appears to constitute the major binding activity at functionally important *cis* elements on the somatostatin promoter. As STF-1 can also stimulate transcription *in vitro* and *in vivo* through the same elements, our data suggest that this homeobox factor is a key regulator of somatostatin expression in small intestine and pancreas.

STF-1 is produced in other pancreatic and intestinal endocrine cell types that do not express detectable levels of somatostatin. STF-1 RNA was distributed throughout the pancreatic islet and small intestine, for example, and was detected in RIN 5AH cells, which produce insulin, but not somatostatin. As STF-1 protein appears to recognize functional elements within the insulin promoter as well (unpublished observations), our results suggest that this factor may have a more general role in promoting endocrine peptide gene expression within the pancreas.

The STF-1 homeodomain is identical to that portion of XIHbox-8 for which sequence is currently available, but sequences C-terminal to that region are unrelated. Although both proteins are expressed in the small intestine, XIHbox-8 protein is apparently restricted to pancreatic acinar cells, whereas STF-1 expression is limited to islet cells. The homology between these two proteins suggests that they may arise as alternate splice products from the same gene or from two genes whose differential expression may determine exocrine *vs.* endocrine fate in the developing pancreas.

We have previously observed that the cell-specific expression of the somatostatin gene requires a CRE sequence in addition to the TSE sites within the somatostatin promoter (3). As CREB constitutes the major CRE-binding activity in Tu-6 cells, we speculate that the cooperativity observed *in vivo* may arise from interactions between CREB and STF-1. In the absence of

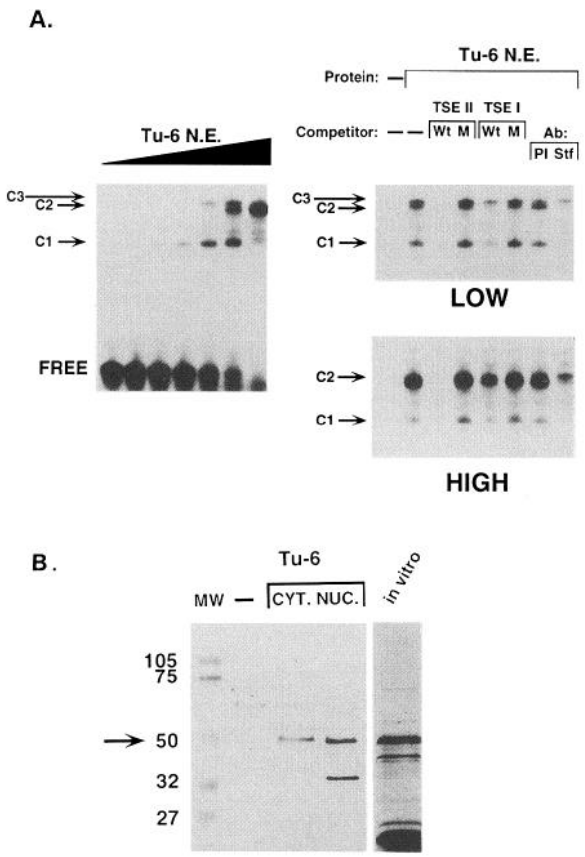

Fig. 6. The Major TSE-Binding Activity in Crude Tu-6 Nuclear Extracts Corresponds to a STF-1-Like Protein

A, Gel mobility shift assay using TSE-II probe on increasing amounts of Tu-6 nuclear extract, as indicated by the *crescendo sign*. *Arrows* point to three protein DNA complexes, named C1, C2, and C3 based on relative mobility. Free, Free probe. *Right*, DNA competition and STF-1 antibody interference studies using low (LOW) or high (HIGH) concentrations of Tu-6 nuclear extract in a gel mobility shift assay. The positions of complexes are indicated. Protein, Presence (Tu-6 NE) or absence (−) of Tu-6 nuclear extract in gel mobility shift reactions. Competitor, Presence (TSE-I and TSE-II) or absence (−) of 20 ng unlabeled wild-type (WT) or mutant (M) TSE oligos in gel mobility shift reactions. Ab, Presence of preimmune (PI) or STF-1 (Stf) antiserum in gel shift reactions. B, Western blot analysis of nuclear (NE) or cytoplasmic (C) extract. MW, Mol wt marker, with relative mass (in kilodaltons) indicated alongside. (*in vitro*), ^{35}S-Labeled *in vitro* translation product derived from rabbit reticulocyte lysate programmed with STF-1 RNA. The *arrow* points to 49-kDa STF-1 product.

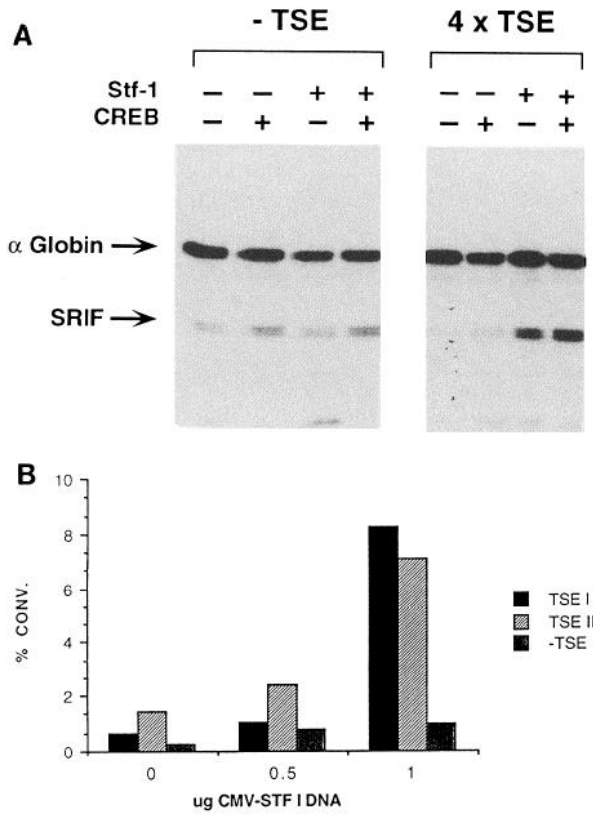

Fig. 7. STF-1 Transactivates the Somatostatin Promoter *in Vitro* and *in Vivo*

A, Primer extension analysis of in vitro transcripiton reactions using HeLa nuclear extracts supplemented with purified recombinant STF-1 and CREB, as indicated (+ or −) *over each lane*. −TSE, Minimal somatostatin promoter construct containing the CRE site, but lacking TSE sites. 4 × TSE, Somatostatin promoter vector containing four TSE-I sites placed up-stream of the CRE site. α-Globin, *Arrow* points to primer-extended product arising from human α-globin template, used as an internal control. SRIF, The *arrow* points to product corresponding to somatostatin templates −TSE or 4 × TSE. B, Representative transient CAT assay of STF-1 activity in HeLa cells. Cells were cotransfected with increasing amounts of CMV-STF-1 expression plasmid, as indicated, plus somatostatin reporter plasmid and Rous sarcoma virus-luciferase internal control vector. TSE-I, Somatostatin-CAT reporter containing four copies of the TSE-I site: TSE-II, somatostatin-CAT reporter with two copies of the TSE-II site; −TSE, somatostatin reporter with no TSE sites; % CONV., Percentage of [^{14}C] chlorampenicol converted to acetylated forms. CAT assays were performed after normalizing to luciferase activity from an internal control plasmid.

hormonal stimulation, CREB activity does not appear to arise from the cAMP-regulated protein kinase-A site, but from a glutamine-rich domain termed Q2 (Montminy, M., submitted). Thus, CREB may subserve several functions by alternatively employing phosphorylation-dependent or constitutive activation domains that synergize with cell-specific factors to provide high level expression of the somatostatin gene. As restricted expression of other genes, including CG and the T-cell receptor, also require a CRE site that synergizes with tissue-specific elements (3), the mechanism we described for somatostatin may be similarly employed for these genes.

MATERIALS AND METHODS

Cell Lines, cDNA Libraries, and Transfections

Tu-6 cells were maintained, passaged, and transfected using calcium phosphate coprecipitation, as previously described

(3). In all cases, CAT activity was measured after normalizing to the activity of a cotransfected Rous sarcoma virus-luciferase reporter plasmid. CMV-STF-1 expression plasmids were constructed by inserting the STF-1 cDNA into a CMV parent plasmid using standard cloning procedures. −TSE and 4 × TSE somatostatin plasmids were previously described (3).

PCR Amplification and Isolation of cDNAs

Homeobox sequences were isolated from a Tu-6 cDNA library by PCR amplification with degenerate oligonucleotides encoding amino acids LEKEF (sense orientation, amino acids 17–21 of the homeodomain) and IWFQN (antisense orientation, amino acids 48–52), which are highly conserved in almost all homeobox proteins. The synthetic primers employed were 5′-GGCGGATCCCTXRARARRGART (A/T) C-3′- and 5′-GGCGGATCCC (G/T) RTTYTGRAACCA-3′, where R = A/G and Y = C/T. PCR was performed using 20 pmol of each primer and 1 ng Tu6 cDNA. The annealing temperature was 45 C for three cycles, followed by 55 C for 35 cycles. The anticipated PCR product of 129 bp was resolved by agarose gel electrophoresis, excised from the gel, subcloned into Bluescript SKII, and analyzed by double stranded DNA sequencing.

To obtain full-length STF-1 cDNA clones, the STF-1 PCR fragment was labeled to high specific activity by random primer labeling. Approximately 10^6 plaques from a Tu-6 λgt11 library were screened by hybridization to the STF-1 fragment probe. Thirty positive plaques were purified, and several of these were subcloned into Bluescript SKII and sequenced on both strands.

RNAse Protection and *in Situ* Hybridization Analysis

RNA was prepared from various cell lines or adult rat tissues by a standard acid-guanidinium-phenol procedure. For RNAse protection analysis, we used a *Hin*dIII-linearized plasmid containing 318 bp of Stf-1 cDNA corresponding to amino acids 60–165. The riboprobe was annealed to 30 μg total RNA and processed as described previously (3).

For *in situ* hybridizations, pancreas and small intestine were sectioned on a cryostat, mounted onto slides, and hybridized with STF-1 antisense riboprobe, as previously described (10).

Tu6 cDNA Library Construction

Tu6 cDNA was synthesized from 5 μg poly(A)-selected RNA using a TimeSaver cDNA synthesis kit (Pharmacia, Piscataway, NJ). *Not*I/*Eco*RI adapters with *Eco*RI overhangs were ligated onto the cDNA ends, and unincorporated linkers were separated from the cDNA by chromatography with a Sepharose CL-4B column (Pharmacia). Complementary DNAs of 1.5–4 kilobases were size-selected by agarose gel electrophoresis and ligated into λgt-11 phage arms. After packaging with Gigapack II Gold (Stratagene, La Jolla, CA), the library contained 4×10^6 plaque-forming units, with less than 2% corresponding to religated phage arms, as determined by blue/white color selection.

Antibodies and Western Blot Analysis

STF-1 antiserum was raised in rabbits using a synthetic STF-1 peptide extending from amino acids 196–214. Western blot analysis with STF-1 antiserum was performed on cytoplasmic and nuclear Tu-6 extracts, as previously described (3).

Expression of STF-1 Protein in *E. coli*

The STF-1-coding sequence was inserted in-frame into the bacterial expression vector pGEX-2T. The resultant plasmid was introduced into the *E. coli* strain BL 21. Cells were grown in 1 liter luria broth medium plus 30 μg/ml ampicillin, to an OD A_{600} of 0.6. Cells were induced with 0.25 g isopropyl thio galactoside for 3 h, spun, and resuspended in HDB buffer (140 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, and 25 mM HEPES, pH 7.4) containing protease inhibitors (1 mM phenylmethylsulfonylfluoride, Trasylol, and 100 U/ml leupeptin). Cells were lysed by treatment with lysozyme (1 mg/ml) for 30 min on ice. Lysis solution was then added such that the extract contained 1% Triton X-100, 5 mM EDTA, 1 mM dithiothreitol (DTT), and 1 M NaCl (final concentrations). The lysate was centrifuged for 30 min at 40,000 × *g*, and the supernatant was dialyzed in HDB buffer containing 1% Triton, protease inhibitors, and DTT at 1 mM for 2 h. The lysate was then mixed with 500 μl glutathione-agarose beads for 20 min at 4 C. The beads were washed seven times, and recombinant STF-1 protein was eluted by incubating the beads with 7 U thrombin for 1 h at room temperature.

Gel Mobility Shift and DNAse-I Protection Analysis

Gel shift assays were performed, as previously described (11), using double stranded somatostatin TSE-I and TSE-II oligos extending from −104 to −86 (5′-TTGCGAGGCTAATGG-TGCG-3) and −303 to −281 (5′-GATCTCAGTAATAAT-CATGCAG-3′), respectively. Mutant TSE-I was described previously (3), and TSE-II oligo contained 5′-GATCT-CAGGCCGGCCGCATGCAC-3′. DNAse-I protection studies with somatostatin promoter fragments were performed as previously described (11).

In Vitro Transcription

The effect of recombinant STF-1 on transcription from the somatostatin promoter was analyzed by *in vitro* transcription assays as previously described (12), except that HeLa nuclear extracts were used in the place of PC12 extracts. Briefly, reactions were carried out in a final volume of 50 μl containing 10 mM HEPES (pH 7.9), 60 mM KCl, 0.2 mM EDTA, 5 mM $MgCl_2$, 5% glycerol, 2% polyvinyl alcohol, 2 mM DTT, 100 ng pUCα1 (control DNA template), 200 ng somatostatin promoter template, 83 μg nuclear extract, and recombinant transcriptional activators, as described in the figure legends. The DNA templates, nuclear extract, and activator proteins were allowed to assemble for 30 min at 30 C before the addition of all four ribonucleotides to final concentrations of 400 μM each. After an additional 30 min of incubation, the reactions were terminated by extraction with phenol-chloroform-isoamyl alcohol (50:49:1) and analyzed by primer extension analysis, as previously described (12). The extension product from the α-globin promoter is 64 nucleotides, and that from the somatostatin promoter is a doublet of 56 and 57 nucleotides.

Acknowledgments

We thank A. Kakizuka for help in cDNA library construction, J. Rivier and W. Vale for synthetic peptide and antiserum preparation, and members of the Montminy lab for helpful discussions.

Received April 28, 1993. Revision received June 17, 1993. Accepted July 13, 1993.

Address requests for reprints to: Dr. Marc R. Montminy, The Clayton Foundation Laboratories for Peptide Biology, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037.

This work was supported by NIH Grants GM-37828 and CA-54418 and by the Foundation for Medical Research, Inc.

* FMR investigator.

REFERENCES

1. Wessells N, Evans J 1968 Ultrastructural studies of early morphogenesis and cytodifferentiation in the embryonic mammalian pancreas. Dev Biol 17:413–446
2. Alpert S, Hanahan D, Teitelman G 1988 Hybrid insulin genes reveal a developmental lineage for pancreatic endocrine cells and imply and relationship with neurons. Cell 53:295–308
3. Leonard J, Serup P, Gonzalez G, Edlund T, Montminy M 1992 The LIM family transcription factor Isl-1 requires cAMP response element binding protein to promote somatostatin expression in pancreatic islet cells. Proc Natl Acad Sci USA 89:6247–6251
4. Vallejo M, Miller C, Habener JF 1992 Somatostatin gene transcription regulated by a bipartite pancreatic islet D-cell-specific enhancer coupled synergistically to a cAMP response element. J Biol Chem 267:12868–12875
5. Ohlsson H, Thor S, Edlund T 1991 Novel insulin promoter- and enhancer-binding proteins that discriminate between pancreatic α and β-cells. Mol Endocrinol 5:897–903
6. Scott M, Tamkun JW, Hartzell GW 1989 The structure of the homeodomain. Biochim Biophys Acta 989:25–48
7. Gehring W 1992 The homeobox in perspective. Trends Biochem Sci 17:277–280
8. Wedeen CJ, Kostriken RG, Matsumura I, Weisblat DA 1990 Evidence for a new family of evoutionarily conserved homeobox genes. Nucleic Acids Res 18:1908
9. Wright C, Schnegelsberg P, DeRobertis EM 1988 XlHbox 8: a novel *Xenopus* homeo protein restricted to a narrow band of endoderm. Development 104:787–794
10. Lee S, Imaki T, Vale W, Rivier CR 1990 Effect of prenatal exposure to ethanol on the activity of hypothalamic-pituitary-adrenal axis of the offspring: impartance of the time of exposure to ethnol and possible modulating mechanisms. Mol Cell Neurosci 1:168–177
11. Yamamoto KK, Gonzalez GA, Menzel P, Rivier J, Montminy MR 1990 Characterization of a bipartite activator domain in transcription factor CREB. Cell 60:611–617
12. Gonzalez GA, Menzel P, Leonard J, Fischer WH, Montminy MR 1991 Characterization of motifs which are critical for activity of the cyclic AMP-responsive transcription factor CREB. Mol Cell Biol 11:1306–1312