Regulation of the human Na\(^+\)-glucose cotransporter gene, SGLT1, by HNF-1 and Sp1

MARTÍN G. MARTÍN, JIAFANG WANG, R. SERGIO SOLORZANO-VARGAS, JASON T. LAM, ERIC TURK, AND ERNEST M. WRIGHT

Departments of Physiology and Pediatrics, Division of Gastroenterology and Nutrition, UCLA School of Medicine, Los Angeles 90095-1751; and Department of Biology, California State University Northridge, Northridge, California 91330

THE SODIUM-GLUCOSE TRANSPORTER SGLT1, an integral membrane protein, is located in the apical membrane of the small intestinal enterocyte and transports glucose and galactose against their concentration gradients (46). The human SGLT1 gene is located on the long arm of chromosome 22 (22q13.1), spans 70 kb in length, and consists of 15 exons (41, 42).

SGLT1 is the prototype of the Na\(^+\)-dependent cotransport family of proteins. Mutations in the coding sequence of its gene are responsible for the autosomal recessive disorder glucose-galactose malabsorption [On line Mendelian Inheritance in Man (OMIM) #182380] (14, 20, 21, 43). Expression of SGLT1 in humans is largely limited to small intestinal enterocytes (12). SGLT1 expression in vivo has been shown to be regulated by various dietary, hormonal, and hard-wired stimuli (7). Intestinal glucose transport capacity in the sheep is particularly dependent on dietary glucose, where SGLT1 protein levels are regulated primarily at a posttranscriptional level (16). In rats, glucose transport was shown to be influenced by a diurnal trigger rather than by changes in dietary glucose (7, 33). In support of this, nuclear run-on assays, performed with nuclei isolated from intestinal epithelial cells, revealed a diurnal pattern of rat SGLT1 transcription that fluctuates by as much as sevenfold (33).

The regulation of the SGLT1 gene provides a model system for evaluation of the mechanism of intestine-specific expression. Compared with our understanding of hepatocyte- and lymphocyte-specific regulation of expression, relatively little is known about trans-acting DNA-binding proteins that affect enterocyte-specific gene expression (6, 40). We do know that expression of the sucrase-isomaltase gene is controlled by cdx-2 and hepatic nuclear factor (HNF)-1\(\alpha\) homo- and heterodimers, and the lactase gene by the homeodomain protein HOXC11 and the GATA family of nuclear proteins (26, 40). However, transgenic mice expressing these critical cis-elements failed to show correct tissue- and age-specific expression (19). Therefore, despite progress, the transcription factors that can direct expression in epithelial cells of the small intestine have not been identified (36).

To gain better insight into the tissue-specific, developmental, and diet-induced regulation of human SGLT1, we have evaluated the constitutive regulation of the gene, determined a minimal promoter, and identified its critical elements using in vitro assays, including reporter analysis of deletion- and substitution-mutant clones, band-shift, DNase I footprint, and overexpression assays.

EXPERIMENTAL PROCEDURES

Cloning of the 5′-flanking region of SGLT1. Cosmid clone H33 contains 25 kb of the 5′-untranslated region (UTR) of SGLT1 (42). H33 was digested with restriction endonucleases, and a Southern blot was probed with an oligonucleotide corresponding to the region immediately upstream of the promoter ([−10/+10] 5′-CTGGCAGAGGGAGGACGC-3′). A 5.3-kb Nco I fragment (−5370/+22) corresponding to the SGLT1 5′-UTR contiguous to the initiator methionine codon...
was subcloned into the luciferase reporter vector pGL3 Basic (Promega) and is referred to as hSGLT1

SGLT1 nested deletion clones. Nested deletions of the SGLT1 5'-UTR were made by exonuclease and mung bean nuclease digestion (22). Clones were isolated, and their sizes were determined by restriction digestion and sequencing. Constructs hSGLT1 3085/-22/B-Luc, hSGLT1 1304/-22/B-Luc, hSGLT1 1040/-22/B-Luc, hSGLT2 1641/-22/B-Luc, and hSGLT2 263/-22/B-Luc were chosen for transient transfection analysis. The transcription factor databases of MacVector 5.0 (TFDSITES.SUBSEQ.7.0.aa) and MatInspector 2.1 (http://www.gsf.de/cgi-bin/matsearch.pl) were searched for putative DNA cis-elements (32).

Chimeric SGLT1 promoter-luciferase constructs were created by cloning PCR-derived fragments of the 5'-flanking region into the vector pGL3 Enhancer. PCR was performed with 11 different sense oligonucleotides and an antisense oligonucleotide corresponding to nucleotides +56/+76 from exon 1 (5'-TGAGATATGGACGACCATTTTCTTTT; and GC box 2/unrelated nucleotides containing +56/+76) were used to amplify the full-length product. These fragments were gel purified, subcloned into the pGL3-Enhancer vector, and sequenced to confirm the mutations.

DNase I footprinting. In vitro DNase I footprint analysis was performed with the hSGLT1 330/-22/E-Luc and hSGLT1 230/-22/E-Luc clones digested with Hind III, radiolabeled with [γ-32P]dATP (6,000 Ci/mmol), and then cut with Nco I (5' to 3') (17).

Band-shift assays. Band-shift analysis was done as previously described, with the exception of the GC boxes that were analyzed in the presence of 1 mM ZnCl2 (17). Standard competition studies were run using excess cold oligonucleotides. The primers used included wild-type (WT) 19–21 primer, spanning from –51 to –30, sense (5'-GACAAGTTGGTGCTGATCATTAACCAGGAGGC), and –300 to –314 (5'-gcaagcttCTGTGGGAGTACAGTGGG-3'). Restriction enzyme and DNA-sequencing analyses.

Scanning mutagenesis. Scanning mutagenesis was done to define the essential constitutive regions of the minimal promoter of SGLT1. Mutagenizing primers (50-mers), containing a central core of 10 mutated nucleotides (A→C, C→A, G→T, T→G) and flanked on either side by a span of 20 correct nucleotides, were designed (Fig. 1A). These mutant oligonucleotides were used with an antisense oligonucleotide from exon 1 (nucleotides +56/+76) to perform the initial PCR amplification. The purified fragments M3 through M23 served as a "megaprimer," which together with the –235/–212 Hind III primer were used to amplify the full-length product. These products were then digested with Hind III and Nco I, gel purified, and subcloned into a similarly digested pGL3-Enhancer vector. Clones hSGLT1 300/-22/E-Luc, hSGLT1 290/-22/E-Luc, hSGLT1 280/-22/E-Luc, hSGLT1 263/-22/E-Luc, hSGLT1 1304/-22/E-Luc, and hSGLT1 1040/-22/E-Luc were all confirmed by restriction enzyme and DNA-sequencing analyses.

RESULTS

The promoter region of human SGLT1 gene. The 5'-untranslated sequence of the human SGLT1 isoform is shown in Fig. 1B and is aligned for direct comparison with the rat (GenBank accession no. 9AF007832) and sheep (GenBank accession no. AJ223077) isoforms (33, 42, 45). Analysis of the 1.5-kb region immediately upstream of the initiator codon for putative cis-acting elements revealed putative binding sites for a single HNF-1 site and a CTC site (5'-CTCGGTCCCTCTCAGTACGTGAG). A mutant (Mut) 19–21 primers were used in combination with antisera provided by G. Crabtree (HNF-1α and HNF-1β, 1 µl each) and Santa Cruz Biotechnology (2 µl of either Sp1, Sp2, or Sp3). Recombinant Sp1 (Promega) was used at a concentration of 0.5 footprint units (FPU) per gel-shift reaction.
morphisms were identified between the two sequences (not shown; GenBank accession no. pending). Transiently transfected Caco-2, LLC-PK1, and CHO cells. Reporter constructs containing shortened lengths of the promoter region were cloned into the luciferase reporter construct pGL3 Basic and were transfected into Caco-2, LLC-PK1, and CHO cells. In Caco-2 cells, the longer clones had promoter activities, 10-fold higher than the promoterless vector, pGL3 Basic (Fig. 2A). Transfection of the shortest clone (hSGLT1 -243/-122) into Caco-2 cells resulted in a 16-fold higher level of promoter activity compared with pGL3 Basic. In contrast, promoter activity was generally lower (eightfold above the empty vector) when tested in LLC-PK1 cells (Fig. 2A). Finally, CHO cells could not support the activity of any size promoter construct. We interpreted these results to suggest that Caco-2 and LLC-PK1 cells are capable of supporting the basal promoter activity of a limited region of the human SGLT1 promoter. These data are consistent with reports that SGLT1 is expressed in Caco-2 and LLC-PK1 cells (18, 29).

To define the lower size limits of the minimal promoter, 11 chimeric SGLT1 promoter-luciferase constructs were transiently transfected into Caco-2 cells. Fragments ranging from the longest (hSGLT1 -5295/-22) to a shorter done (hSGLT1 -239/-22) resulted in similar promoter activity, approximately sixfold higher than the promoterless control vector (Fig. 2B). In contrast, promoter activity of shorter clones was indistinguishable from that of the control. Deletion analyses of various lengths (-5295/-22 to -27/-22) of the SGLT1
promoter suggest that its expression was supported best in the human intestinal cell line Caco-2 and that nucleotides -235/+22 represent the gene’s minimal promoter (Fig. 2).

DNase I footprint analysis identifies two DNA-protein complexes. Using 5’ to 3’-labeled hSGLT1 -330/+22 as a template revealed a single DNA-protein complex in the region corresponding to -229 through -206 (named FP-I; Fig. 3). Similarly, a DNase I digest of the hSGLT1 -235/+22 vector labeled in the same orientation revealed a second complex located at -49 to -31 in the upstream region of the gene (named FP-II; Fig. 3). No other footprints were identified within bases -270 to +30.

Fig. 2. Transiently transfected Caco-2 and LLC-PK1 cells support promoter activity of various nested deletion clones. A: relative size of the upstream region of each clone is shown. Calcium phosphate methods were used to cotransfect 1 pmol of vectors of various sizes and 0.1 pmol of cytomegalovirus (CMV)-β-galactosidase vector. Transfection efficiency and basal luciferase expression were controlled for by transfection of pGL3 Control (containing SV40 enhancer and promoter) and pGL3 Basic. Samples were processed 2 days later, and relative luciferase units (RLU)/β-galactosidase activity are displayed as multiples of elevation over promoterless pGL3-Basic vector. Values are means ± SD of triplicate data from 2–4 experiments. CHO, Chinese hamster ovary. B: deletion fragments subcloned into the pGL3-Enhancer vector (1 pmol). Values are means ± 5D of triplicate data from 2–4 experiments.
HNF-1 is a potent activator of the SGLT1 minimal promoter. The nucleotide sequence from −51 to −30 accords with a transcription factor HNF-1 consensus (5′-GGTTAAATATTAAACCa/c-3′). This domain, corresponding to the M20 clone and footprint II, was further evaluated by band-shift analysis. The WT 19–21 primer (nucleotides −51 to −30) was radiolabeled and found to compete with the Mut 19–21 primer, which contains a critical 4-bp mutation in the HNF-1 site. Figure 5 is a representative band-shift study using the HNF-1 probe that reveals a DNA/protein complex that runs as a smear. This complex was specific because it was competed with a 10-fold excess of the unlabeled HNF-1 oligonucleotide (Fig. 5) but not with either the Mut 19–21 or an unrelated oligonucleotide. Supershift experiments showed a shift of a portion of the specific complex with antiserum to the HNF-1-α isoform (Fig. 5) and its complete removal by antiserum to the HNF-1-β isoform (Fig. 5).

Either the minimal promoter construct (hSGLT1−235/+22/E-Luc) or the corresponding promoter construct containing a 10-bp mutation (M20) in the putative HNF-1 site was used to transiently transfect Caco-2 cells, which were cotransfected with vectors expressing the murine versions of either HNF-1-α or β. Cotransflecting with the HNF-1-α expression vector and the minimal promoter construct resulted in a nearly threefold enhancement of promoter activity over WT alone (Fig. 6). Similarly, a modest twofold increase in promoter activity was obtained using the HNF-1-β expression vector. The action of the HNF-1 expression vector on promoter activity was specific to the site identified at −45/−36 because cotransfection with the M20 mutant substituting for the minimal promoter construct abolished inducible transcriptional activity.

Two regions form specific DNA-protein complexes. A double-stranded oligonucleotide named WT 1–3 was made that encoded domains M1 to M3 (see Fig. 1B; bases −235 to −206). A double-stranded oligonucleotide named WT 17–18 was developed that encodes for domains M17 and M18 (see Fig. 1B; bases −75 to −56).

The WT 1–3 probe revealed two prominent DNA-protein complexes (arrowheads) that could be competed entirely with the addition of 90-fold excess of cold WT 1–3 oligo (Fig. 7). Two DNA-protein complexes were also seen with the WT 17–18 probe, but this complex could not be competed with as much as 250-fold excess of cold WT 17–18 duplex (Fig. 7).

GC boxes 1 and 2 bind recombinant Sp1. Because the WT 1–3 and WT 17–18 duplexes share DNA-protein complexes of similar size and sequence (gtCCTCCT-CCC), we hypothesized that the complexes may be attributed to binding of the transcription factor Sp1 to the GC box. To compare the ability of the WT 1–3 (GC box 1) and WT 17–18 (GC box 2) duplexes to bind recombinant Sp1, the WT 1–3 primer was labeled and competition studies were performed with various multiples of excesses of either itself (Fig. 8) or the WT 17–18 duplexes (Fig. 8). Recombinant Sp1 and the WT 1–3 primer form a prominent complex that corresponds...
in size to that of the slower migrating complex seen with crude nuclear extracts (Fig. 8). Overexposed autoradiograms revealed an additional complex whose molecular weight was approximately twice that of the main complex and may represent Sp1 dimers (Fig. 8). Competition with a 90-fold excess of either the consensus Sp1 or WT 1–3 duplexes was more effective than with WT 17–18 (Fig. 8). In addition, the CTC duplexes failed to compete (Fig. 8). Recombinant Sp1 formed a complex with the labeled WT 17–18 fragment, and this complex could be competed with a 240-fold excess of cold probe (Fig. 8). The slower migrating complex seen with Sp1 and WT 1–3 was not visualized in even the overexposed autoradiogram of the WT 17–18 probe (Fig. 8). Overall, these data were interpreted to suggest that the WT 1–3 (GC box 1) duplexes were capable of binding to Sp1 as efficiently as a consensus Sp1 duplexes and significantly better than the WT 17–18 (GC box 2) duplex.

Sp1, Sp2, and Sp3 bind to GC boxes 1 and 2. To determine whether other members of the Sp1 could account for the DNA-protein complex, supershift assays were performed. A representative gel of GC box 1 demonstrates that the lower portion of the slower migrating complex is shifted by the Sp1 antiserum (Fig. 9). Similarly, Sp2 antiserum shifted a complex that originated in a location that was similar to Sp1 (Fig. 9). In contrast, Sp3 antiserum shifted the entire faster migrating complex and a portion of the slower complex when added with the Sp1 antiserum (Fig. 9). Finally, the addition of all three antisera resulted in residual complex formation that is very similar to what was seen with the addition of both the Sp1 and Sp3 antisera (Fig. 9). These supershifted complexes were specific since the same quantity of preimmune rabbit serum failed to produce a similar supershifted complex (Fig. 9).

GC box 1 is flanked downstream by nucleotides critical for binding recombinant Sp1. To determine which nucleotides of the WT 1–3 oligonucleotide are necessary for binding Sp1, band-shift assays were performed with a 500-fold excess of a series of cold oligonucleotide duplexes containing specific mutations (Fig. 10A). Each competing oligonucleotide contained a three contiguous base pair mutation, the exact location of which is displayed in Fig. 10A. Figure 10B shows a band-shift study with labeled WT 1–3 oligonucleotide and 7 µg of Caco-2 nuclear extracts. Two complexes were identified that could be competed with 500-fold excess of a series of cold oligonucleotide duplexes containing specific mutations (Fig. 10A). Each competing oligonucleotide contained a three contiguous base pair mutation, the exact location of which is displayed in Fig. 10A. Figure 10B shows a band-shift study with labeled WT 1–3 oligonucleotide and 7 µg of Caco-2 nuclear extracts. Two complexes were identified that could be competed with 500-fold excess of the WT 1–3 primer (Fig. 10B). Primers that contained mutations in regions a, b, l, and m were capable of binding the protein(s) that account for the two complexes (Fig. 10A). Moreover, it appears that mutant primers g–k were able to
Binding of Sp1 to the WT 17–18 primer is limited to GC box 2. Similar studies were performed with GC box 2 (the WT 17–18 oligonucleotide) and crude Caco-2 extracts. Figure 11B demonstrates that primers containing mutations in regions a–d and h could compete for binding of the nuclear proteins that are responsible for complex formation. Furthermore, primers with mutations in the e–g region (Fig. 11B) failed to compete, suggesting that nucleotides −51 to −59 (boxed nucleotides in Fig. 11A) are critical for the formation of the DNA-protein complex. The WT 1–3 primer could compete for binding of the complex formed with the WT 17–18 primer (Fig. 11B). Additional evidence suggesting that the complex is related to binding of Sp1 is provided by the ability of consensus Sp1 primer to compete for the two complexes (Fig. 11B). Similarly, the CTC oligonucleotide was able to partially compete for binding (Fig. 11B).

Proteins of the Sp1 multigene family and HNF-1 synergistically activate transcriptional activity of SGLT1. To determine whether the identified HNF-1 and GC boxes influence the activity of a heterologous promoter, we subcloned these elements just upstream of the SV40 promoter in the reporter vector pGL3 Promoter. All constructs contain a single copy of the identified element oriented in the 5′-to-3′ direction. Because of the close apposition of the downstream GC box 2 (−63 to −55) to the HNF-1 element (−51 to −37), we also tested whether or not the two elements altered promoter activity in a synergistic manner. Compared with the enhancerless SV40 promoter construct, the addition of the GC box 1 (WT 1–3) failed to alter promoter activity (Fig. 12). Similarly, neither the HNF-1 site nor GC box 2 (WT 17–18) influenced luciferase activity.

We investigated further the nature of the DNA-protein complexes by performing gel-shift assay with recombinant Sp1 (Fig. 10C). In this experiment, a single prominent complex was formed that resembled the slower migrating complex identified with the crude nuclear extracts (Fig. 10C). However, as in Fig. 8, a fainter and more slowly migrating complex was also seen. The addition of 500-fold excess of cold mutant primers a, b, l, and m was capable of entirely competing for recombinant Sp1 (Fig. 10C). Competition with the other primers (c–k) failed to entirely compete for the Sp1 complex, as shown previously in Fig. 10B. Finally, primers (g–k) that contained mutations outside of the critical GCCCCTCCCC region (Fig. 10A), were capable of only partially competing for the formation of the complex (Fig. 10C). Together, these results were interpreted to suggest that Sp1 and other unidentified proteins are capable of forming specific complexes with the GC box and the downstream DNA element (ATTGC-CAGGACAGCTC) located between nucleotides −223 and −208 of the SGLT1 promoter.

**Fig. 5.** Band-shift assays of HNF-1 element identified a specific complex. HNF-1 primer spans from −51 to −30 (WT 19–21). Annealed WT 19–21 primers were labeled with 32P and used for all experiments. Competition experiments were performed with 10-, 50-, 100-, 1,000-fold (lanes 2–5, respectively) excess of cold WT 19–21 primers. In addition, 1,000-fold excess of oligo containing a mutation in HNF-1 site (Mut 19–21) (lane 6) and unrelated (UR) (lane 7) primers were also used. Single DNA-protein complex is indicated with filled arrowheads. Supershift studies were also performed by preincubating nuclear extract with 1 µl of either HNF-1 antisera (lane 8) or HNF-1β antisera (lane 10), and supershifted band is indicated with an open arrowhead.

**Fig. 6.** Transient cotransfection of clones containing either SGLT1 minimal promoter (SGLT1) or a mutation in HNF-1 region (M20 SGLT1) and an expression vector expressing either HNF-1α or β. Cotransfection of 1 pmol of either HNF-1α or β mammalian expression vector with either hSGLT1−235/E-Luc or hSGLT1−235/E-LucM20/E-Luc clone was performed using standard CaPO4 method. Cells were processed 2 days later, and β-galactosidase and RLU were measured. Data are displayed as multiples of increases over promoterless pGL3-E enhancer clone. Values are means ± SD of triplicate data from 2 experiments.

G597
production above what was seen with the enhancerless SV40 promoter vector. However, the addition of both the GC box 2 and the HNF-1 element enhanced promoter activity by 60% compared with control, suggesting that proteins binding to both of these sites synergistically influenced transcriptional activity in the context of the heterologous SV40 promoter.

DISCUSSION

In this study we analyzed the basal pattern of expression of the promoter region of the human SGLT1 gene in the human intestinal cell line Caco-2. Deletion analysis of various lengths (−5295/+2 to −27/+2) of the promoter inserted upstream of the luciferase reporter indicated that nucleotides −235/+22 represent the gene's minimal promoter (Fig. 2). Within this region of the promoter, several DNA-protein complexes were identified by in vitro footprinting (Fig. 3), and mutagenesis identified three distinct sites that were responsible for enhancing the promoter activity of SGLT1 (Fig. 4).

We determined that an HNF-1 element and two GC boxes control SGLT1 basal expression in the Caco-2 cell. Two active cis-elements were found that bind members of the Sp1 family of proteins within the minimal promoter of SGLT1, and scanning mutagenesis revealed that both elements function to enhance the gene's basal expression (Fig. 4). The components of the complexes were defined by supershift analysis and revealed that portions of the two complexes are formed by binding of Sp1, Sp2, and Sp3 or a related protein(s) (Fig. 9). More specifically, the fastest-migrating DNA-protein complex was completely shifted by Sp3 antisera, whereas the slower-migrating complex appears to be at a minimum a closely migrating triplet, composed of Sp1, Sp2, Sp3, and another undefined complex. Sp1, Sp3, and Sp4 have highly conserved zinc finger DNA binding domains and recognize the consensus GC box (5′-KRGGMGKRRY) with similar specificity and affinities, whereas Sp2 binds with much lower affinity (9). Moreover, whereas Sp1 and Sp4 have only been implicated as transcriptional activators, Sp3 usually functions as a transcriptional repressor and on occasion as an activator (1, 8, 10, 11, 13). Although Sp1 is ubiquitously expressed in all cell types examined, its level of expression varies by as much as 100-fold, and this variability has been implicated in specifying tissue.
and developmental-specific regulation of several genes (15, 37). Similarly, other members of the zinc finger Sp1 multigene family, Sp2 and Sp3, are ubiquitously expressed, whereas Sp4 expression is limited to the brain (9). However, an exhaustive analysis of the tissue distribution, particularly in the intestine, has not been performed for any member of the growing family of Sp1-like proteins (37). Because of the disparate levels of expression, affinity, and function of the Sp1 multigene family members, the overall impact that they may have on SGLT1 expression may be dramatic and is currently under investigation.

Gel-shift assays of GC box 1 revealed that nucleotides immediately downstream of the box (5'-ATTCG-CAGGACAGCTC) were critical for complex formation with both crude and recombinant Sp1 protein (Figs. 10 and 13). Interestingly, although GC box 1 is conserved in both rat and sheep promoters, this adjacent sequence is not conserved and does not resemble a GC box (Fig. 13). Most good Sp1 binding sites are 10 nucleotides in length and do not differ from the consensus sequence at more than one position. Furthermore, Sp1 is only able to bind simultaneously to adjacent Sp1 sites if the central portions of the elements are more than 10 nucleotides apart (2). Since Sp1 is clearly capable of binding to the GC box 1, we would expect that if two Sp1 sites are occupied simultaneously, only the furthest downstream portion of the sequence would be able to bind Sp1. In fact, we failed to identify evidence that two Sp1 monomers could bind simultaneously to the GC box 1 and the adjacent downstream sequence. The faint and slowly migrating complex seen with recombinant Sp1 (Figs. 8 and 10) probably represents the multimerized form of Sp1, a process that can occur at high protein concentrations (2, 27). Therefore, it would be rather surprising if this downstream sequence could bind Sp1 independently of GC box 1. What remains unclear is how Sp1 and related proteins are capable of interacting with the adjacent nucleotides, since they do not resemble a GC box (Fig. 13). It is conceivable that the specificity of binding, including the process of multimerization, may be governed by the nucleotides that flank GC box 1.

Can the Sp1 family of proteins that bind to the two GC boxes interact with one another to synergistically activate expression of the SGLT1 gene? In many genes, GC boxes frequently occur as multiple repeat sequences (2). The repeat Sp1 sites may be either adjacent to one another or widely separated and at either distance are capable of inducing synergistic transactivation (2). Sites that are at close proximity may undergo protein-protein interactions that influence cooperative binding, whereas distal GC boxes may form multimeric complexes that enhance DNA binding and bring elements closer to the core transcriptional machinery (2). This form of synergistic activation between distal and proximal GC boxes may result from self-association of Sp1 and enhanced activity of the transcriptional complex.

Scanning mutagenesis identified that the putative HNF-1 element was also active in inducing SGLT1 basal expression. HNF-1 induces the expression of several intestinal genes, including sucrase-isomaltase, aminopeptidase, apo B, lactase, α-fetoprotein, α1-antitrypsin, and aminopeptidase N (26, 40). HNF-1α is expressed in the small intestine, kidney, stomach, and liver, whereas HNF-1β is produced in the ovary, lung, and small intestine; expression of both proteins is limited to villus epithelial cells (23, 38). The relative concentrations of HNF-1α and -β differ markedly from tissue to tissue and may be developmentally regulated (44).

Rhoads et al. (33) have implicated HNF-1 in altered SGLT1 expression during the normal day and night cycles of rodents. Nuclear run-on experiments in rat intestinal epithelial cells showed that this regulation occurs at the level of transcription. Moreover, bandshift assay with nuclear extracts isolated from rat intestine identified a HNF-1 protein-DNA complex that migrated as a smear in the evening and as a faster
Fig. 10. Band-shift assay defines nucleotides in the WT 1–3 region that binds Caco-2 nuclear extracts. A: competing primers were used in 500-fold excess, and sequence of sense strand is displayed. Location of 3-bp mutation is displayed in bold, italicized, lowercase letters, and name of each primer is displayed just to left of primer. Nucleotides that are essential for binding nuclear protein are depicted within larger box in WT primer. Location of GCCCTCCCC sequence is outlined. All primers were annealed to an antisense counterpart (not shown). B: double-stranded WT 1–3 primers were labeled with $^{32}$P and incubated with Caco-2 nuclear extracts in presence or absence of a series of cold primers (500-fold excess) that contained mutations. Cold primers used for each competition reaction are indicated above autoradiogram. C: recombinant Sp1 (0.5 FPU/lane) was incubated with WT 1–3 primers in presence or absence of 500-fold excess of cold primer. A multimerized form of Sp1 is shown with an open arrowhead. All primers were annealed to an antisense counterpart (sequence not shown). DNA-protein complexes are shown with closed arrowheads.
Antiserum to the HNF-1α isoform supershifted all complexes, implicating HNF-1α at all time intervals. However, antiserum for the β isoform abolished the migration of primarily the evening complex, indicating that HNF-1β was a component of the evening complex. The authors suggested that this effect could be explained if the epitope recognized by the antibody is part of the HNF-1 DNA binding site. However, previous reports showed that the identical β-antiserum did not disrupt complex formation but rather supershifted HNF-1β when tested by band-shift technique (24). Our data showed that the HNF-1β antiserum also abolished the slower migrating complex (Fig. 5). However, other data from our laboratory suggested that this effect was nonspecific, as other unrelated DNA-protein complexes were also disrupted by this antiserum (data not shown). Thus these data underscore the difficulty of assessing the contribution of the β-isoform by supershift assay using currently available antisera.

The critical role of HNF-1 in controlling the regulation of genes in vivo was clearly shown in HNF-1 knockout mice. These mice developed profound multorgan effects, including a dramatic decline in complex in the morning. Antiserum to the HNF-1α isoform supershifted all complexes, implicating HNF-1α at all time intervals. However, antiserum for the β isoform abolished the migration of primarily the evening complex, indicating that HNF-1β was a component of the evening complex. The authors suggested that this effect could be explained if the epitope recognized by the antibody is part of the HNF-1 DNA binding site. However, previous reports showed that the identical β-antiserum did not disrupt complex formation but rather supershifted HNF-1β when tested by band-shift technique (24). Our data showed that the HNF-1β antiserum also abolished the slower migrating complex (Fig. 5). However, other data from our laboratory suggested that this effect was nonspecific, as other unrelated DNA-protein complexes were also disrupted by this antiserum (data not shown). Thus these data underscore the difficulty of assessing the contribution of the β-isoform by supershift assay using currently available antisera.

The critical role of HNF-1 in controlling the regulation of genes in vivo was clearly shown in HNF-1 knockout mice. These mice developed profound multorgan effects, including a dramatic decline in complex in the morning. Antiserum to the HNF-1α isoform supershifted all complexes, implicating HNF-1α at all time intervals. However, antiserum for the β isoform abolished the migration of primarily the evening complex, indicating that HNF-1β was a component of the evening complex. The authors suggested that this effect could be explained if the epitope recognized by the antibody is part of the HNF-1 DNA binding site. However, previous reports showed that the identical β-antiserum did not disrupt complex formation but rather supershifted HNF-1β when tested by band-shift technique (24). Our data showed that the HNF-1β antiserum also abolished the slower migrating complex (Fig. 5). However, other data from our laboratory suggested that this effect was nonspecific, as other unrelated DNA-protein complexes were also disrupted by this antiserum (data not shown). Thus these data underscore the difficulty of assessing the contribution of the β-isoform by supershift assay using currently available antisera.

The critical role of HNF-1 in controlling the regulation of genes in vivo was clearly shown in HNF-1 knockout mice. These mice developed profound multorgan effects, including a dramatic decline in
survival after weaning (30, 31). They also experience failure to thrive, dramatic hepatic enlargement, severe phenylketonuria, and Fanconi syndrome, including severe glucosuria. The HNF-1 knockout mice had reduced phlorizin-binding to renal brush-border membranes, suggesting a decline in the yet-to-be-identified renal high-affinity/low-capacity cotransporter (SGLT2) (46). Interestingly, the authors did not describe diarrheal symptoms in these mice. If HNF-1 is critical in controlling SGLT1 expression, one would have expected evidence of glucose/galactose malabsorption (and consequent diarrhea) on a lactose-based diet (breast milk) (21). Similarly, the HNF-1β knockout mice have been shown to develop a form of Laron dwarfism and non-insulin-dependent diabetes (5). In humans, heterozygous germline mutations (autosomal dominant) of either HNF-1α or β result in a poorly defined form of diabetes whose onset begins in late adolescence (MODY3) (47). Diarrhea and glucose malabsorption have not been reported in patients with MODY3.

Functional analysis suggested that proteins that bind to GC box 2 and the HNF-1 element synergistically enhanced SGLT1 promoter activity (Fig. 12). Numerous transcription factors have been shown to interact with Sp1, including GATA1, gut-enriched kruppel-like factor, AP1, NF-κB, GATA, and HNF-4 (35, 48). Although Sp1 enhances HNF-4-induced transcription of the apoCII promoter, the nature of the synergy between these factors was not actually defined (39). Analysis of other transcription factors suggests that the Sp1 COOH-terminal domain is involved in both synergistic activation and protein-protein interaction (39). HNF-1, on the other hand, has been shown to interact only with C/EBPα and to synergistically activate expression of the human albumin promoter (25). A serine-threonine- and proline-glutamine-rich region of HNF-1β is critical for its interactions with C/EBPα and its functional synergy at the albumin promoter. Although transcriptional synergism between Sp1 and HNF-1 has never been reported, this type of interaction may explain how a ubiquitous transcription factor like Sp1 may direct intestinal-specific expression. Interestingly, both HNF-1 and Sp1/S3 have been implicated in mediating glucose activation of various promoters (3, 4, 34).

Overall, this study represents the first detailed analysis of the SGLT1 promoter and the critical role that HNF-1 and Sp1 have in controlling basal transcription. In the analysis of the promoter, we have identified several unique features that deserve further attention. Although Sp1, Sp2, and Sp3 influence the basal expression of the gene via two separate GC boxes, the precise role that these and other members of the Sp1 multigene family have in altering expression of the gene has not been determined. Sp1 was shown to interact with GC box 1, but the nature of how the 16-nucleotide element immediately downstream of the box confers specific binding to Sp1 remains unexplored. Finally, the mechanism by which Sp1 and HNF-1 synergistically influence SGLT1 gene expression should be further analyzed since it may be a general model for understanding the mechanism of intestinal-specific gene expression.

This research was supported grants from the National Institutes of Health (HD-34706 and DK-44582), the Robert Wood Johnson Foundation Faculty Training program, and the American Gastroenterology Industry Training Award.

Received 2 November 1999; accepted in final form 9 December 1999.

REFERENCES


