Members of the Caudal Family of Homeodomain Proteins Repress Transcription from the Human Apolipoprotein B Promoter in Intestinal Cells*

(Received for publication, August 29, 1995, and in revised form, October 31, 1995)

Soon-Youl Lee‡, Brian P. Nagyi‡, Alan R. Brooks‡, Duen-Mei Wang‡‡, Bernhard Paulweber††, and Beatriz Levy-Wilson‡‡

From the ‡Palo Alto Medical Foundation Research Institute, Palo Alto, California 94301, ††Gladstone Institute for Cardiovascular Disease, and ‡‡First Department of Internal Medicine, Landeskrankenanstalten, Salzburg, Austria

Apolipoprotein B (apoB) is the major protein component of low density lipoproteins, and plays a central role in cholesterol transport and metabolism. The apoB gene is transcribed in the liver and in the intestine in humans. Although much is known about the DNA sequence elements and protein factors that are important for transcription of the human apolipoprotein B gene in the liver, less is known about the mechanisms that control transcription of this gene in the intestine. The sucrose isomaltase gene (SI), is expressed exclusively in the intestine. Two sequences from the promoter region of the SI gene, namely SIF-1 and SIF-3, are essential for promoter activity of the SI gene in intestinal cells. Sequences displaying a high degree of similarity to those of SIF-1 and SIF-3 are present in the third intron of the apoB gene. Rather than stimulating apoB promoter activity, the BSIF-1 and BSIF-3 sequences repressed transcription in CaCo-2 cells. Gel retardation studies demonstrated that BSIF-1, like SIF-1, binds to proteins related to the caudal family of proteins such as mCdx-4 and mCdx-2. These proteins appear to repress transcription from the apoB promoter by a mechanism that involves an interaction with members of the C/EBP family of proteins, that bind to a target sequence for the repressor in the segment from −139 to −111 of the apoB promoter. On the other hand, BSIF-3, like SIF-3, binds to HNF-1 and also represses transcription from the apoB promoter.

In recent years, a large body of data has been gathered regarding the mechanisms that operate to control the expression of eucaryotic genes in specific tissues and cell types. One such tissue that has received a lot of attention is the liver. Work in many laboratories studying liver-specific gene expression has revealed that a relatively small number of hepatic transcription factors, acting cooperatively with other ubiquitous transcription factors or synergistically with each other influence transcription of many hepatic-specific genes in cultured cell lines and also in transgenic animals (1–4). Many key liver-specific transcription factors have been identified and their mode of binding and influencing transcription has been characterized for some genes. Thus, HNF-1 (5), C/EBP (6), HNF-3 (7) and HNF-4 (8), alone or in combination, appear to play a role in the regulation of transcription of most hepatic-specific genes that have been examined so far, either by binding to the promoter region itself, or to tissue-specific enhancer elements localized upstream or downstream of the transcriptional start site.

Apolipoprotein B is the sole protein component of low density lipoprotein and plays an important role in cholesterol homeostasis and coronary artery disease in humans (9). It is encoded by a single-copy gene in the human genome that extends over 48 kilobases and is transcribed mainly in the liver and intestine (10). Work in our laboratory and in others has defined the cis-acting DNA sequences and the trans-acting protein factors involved in the transcriptional control of the apoB gene in hepatic cells. For example, the hepatic-specific control elements are found in the proximal promoter region between −260 and +1 (11), include binding sites for C/EBP, HNF-3, and HNF-4, and appear sufficient for promoter activity in cultured hepatic cell lines, such as HepG2 cells. Transcription from the apoB promoter in vivo in the liver of transgenic mice requires an enhancer from the second intron of the gene (12) that functions by a mechanism involving synergistic action of three hepatic-enriched transcription factors, namely HNF-1, C/EBP, and protein II (13).

In contrast, very little is known about the DNA sequence elements of the human apoB gene and of the transcription factors that modulate transcription of this gene in intestinal cells. Studies by us and others using cultured colon carcinoma (CaCo-2) cells revealed that 139 base pairs of upstream sequence are sufficient for strong activity of the apoB promoter in these cells. Nevertheless, constructs containing this portion of the apoB promoter linked to a reporter gene fail to be expressed in the intestines of transgenic mice.2 Indeed, even larger constructs, encompassing several kilobases of 5' and 3' apoB DNA sequences that are expressed at high levels in the liver, fail to be transcribed in the intestines of transgenic mice (for review, see Ref. 4) making the search for the intestinal control ele-
ments of the human apoB gene an interesting challenge.

Recent studies by Traber et al. (14) with the sucrose isomaltsa
(14) with the sucrose isomaltsa gene, that is transcribed exclusively in intestinal cells, revealed that two regulatory sequences in the proximal promoter region played an important functional role in transcriptional activation of the SI gene in CaCo-2 cells. These two sequences were designated SIF-1 and SIF-3 (14). Further work demonstrated that the SIF-1-binding protein is a member of the caudal family of homeodomain proteins and is equivalent to mouse mCdx-2 (15). Additionally, SIF-3 was shown to bind HNF-1 (15).

In our quest to identify the apoB sequence elements required for its expression in intestinal cells, we compared DNA sequences of regulatory segments of the human apoB gene and the binding sites for SIF-1 and SIF-3. To our surprise, we identified two sequences within the third intron of the human apoB gene exhibiting high homology to SIF-1 and SIF-3, and we designated them BSIF-1 and BSIF-3. We then examined the functional properties of BSIF-1 and BSIF-3 in apoB promoter activity as well as the binding of tissue-specific transcription factors to those sequences. Our results revealed that BSIF-1 and BSIF-3 repress promoter activity in intestine-derived CaCo-2 cells. Furthermore, BSIF-1, like SIF-1, binds to those sequences of the rat liver and BSIF-3 to that of intestine cells. Both sequences from the apoB third intron repress rather than stimulate promoter activity in intestine-derived CaCo-2 cells. Repression by BSIF-3 appears to involve an interaction with basal components of the transcriptional machinery. On the other hand, repression by BSIF-1 appears to involve an interaction with sequences from −139 to −111 of the apoB promoter.

**EXPERIMENTAL PROCEDURES**

Plasmid Construction—Construction of plasmids, −898 CAT, −139 CAT, −111 CAT, −85 CAT, and TATA-CAT have been described elsewhere (11, 16). Plasmid CEBP-TATA-CAT was made by placing six copies of the C/EBP-binding site from the apoB promoter (−53 to −72), into the Xbal site upstream of the TATA fragment in the TATA-CAT vector. The BSIF-1 oligonucleotide corresponds to the segment from 1952 to 1979 of the apoB gene and the BSIF-3 oligonucleotide corresponds to the sequence from 2358 to 2384 of the apoB gene. Double-stranded BSIF-1 and BSIF-3 oligonucleotides were exposed to T4 polynucleotide kinase and then ligated at 14 °C for 2 h. These oligonucleotides were inserted into either the Xbal site of the −898 CAT, −139 CAT, −111 CAT, and −85 CAT plasmids to make BSIF-1 or BSIF-3 −898 CAT, −139 CAT, −111 CAT, and −85 CAT plasmids, respectively, or into the HindIII site of the −30 CAT and CEBP-TATA-CAT plasmids. Oligonucleotides −139 to −111 and −111 to −85 were then inserted into the Xbal site of the TATA-CAT plasmid. The sequences of all the constructs were verified by DNA sequencing by the dideoxy termination method using Sequanase.

**RESULTS**

**BSIF-1 and BSIF-3 Repress Activity of the Human ApoB Promoter in CaCo-2 Cells**—The sequences of BSIF-1 and BSIF-3 are shown in Fig. 1 together with those of human and mouse SIF-1 and SIF-3. There is a high degree of similarity between the sequences from the sucrose isomaltase promoter and those in the apoB third intron. BSIF-1 is identical at 16 of 24 base pairs (67% identity) with human SIF-1, while BSIF-3 is identical at 13 out of 19 base pairs (68% identity) with human SIF-3. Because SIF-1 and SIF-3 can specifically activate the sucrose isomaltase promoter in intestinal cells, we asked whether the homologous sequences BSIF-1 and BSIF-3 from the apoB third intron may activate the apoB promoter in intestine-derived CaCo-2 cells.

To this end, oligonucleotides corresponding to BSIF-1 and BSIF-3 sequences were cloned upstream of a number of apoB promoter-CAT constructs of varying strengths, and the effects of BSIF-1 and BSIF-3 sequences upon apoB promoter activity were determined in transient transfection experiments. CaCo-2 cells were used as a model of intestinal-derived cells; HepG2 cells represented hepatic derived cells. A segment containing apoB sequences from −898 to +121 (−898 CAT construct) exhibits promoter activity in transient assays with HepG2 and CaCo-2 cells (11). Using this promoter construct, we observed that addition of three copies of BSIF-1 or BSIF-3 upstream of this promoter reduced CAT activity both in CaCo-2 and HepG2 cells by 20–50%, suggesting that a weak negative element may reside in the sequences represented by the BSIF-1 and BSIF-3 oligonucleotides (Fig. 2).

To define the target for this repression by BSIF-1 and BSIF-3 seen with the entire apoB promoter segment, we cloned the BSIF-1 and BSIF-3 oligonucleotides upstream of shorter apoB promoter-CAT constructs and evaluated the impact of these added sequences upon apoB promoter activity in these cell lines (Fig. 2). We first examined the effects of BSIF-1 and BSIF-3 upon the promoter construct extending from −139 to +121 (−139 CAT). This construct displays a strong promoter activity.
in CaCo-2 cells (8-fold higher than that of the −898 CAT construct) and a similar promoter activity in HepG2 cells as compared to −898 CAT (11). Two series of constructs were analyzed; those incorporating one copy of the BSIF-1 or BSIF-3 oligonucleotides upstream of −139 CAT and those incorporating five copies of these oligonucleotides upstream of −139 CAT.

In CaCo-2 cells, a 5-fold repression of the −139 CAT promoter construct was observed when only one copy of either BSIF-1 or BSIF-3 was incorporated upstream of the promoter (Fig. 2), thus suggesting that the BSIF-1 and BSIF-3 oligonucleotides may harbor binding sites for a negative regulator of apoB transcriptional activity. In HepG2 cells, a 2-fold decrease in promoter activity was evident (Fig. 2). When five copies of BSIF-1 or BSIF-3 were present, the reduction in promoter activity was even higher (−12-fold) in CaCo-2 cells, but about the same (−2-fold) in HepG2 cells. These results suggested that BSIF-1 and BSIF-3 exert a powerful negative effect upon apoB transcription in CaCo-2 cells, and a minor (2-fold) negative effect upon apoB promoter activity in HepG2 cells.

To more accurately pinpoint the possible target for this strong repression by BSIF-1 and BSIF-3 in CaCo-2 cells, three additional promoter−CAT constructs were tested. The first construct, −111 CAT, extends from −111 to +121 and exhibits a weaker promoter activity than −139 CAT in HepG2 and CaCo-2 cells (2−3-fold). The same is true for −85 CAT (−85 to +121 CAT) and TATA-CAT (−30 to +121 CAT) constructs, that are even weaker, respectively, than −111 CAT (11). BSIF-1 and BSIF-3 exerted a 2−3-fold negative effect upon the activity of the −111 CAT construct in CaCo-2 cells (Fig. 2). In HepG2 cells, BSIF-1 had a minor negative effect (15%), while BSIF-3 appeared to enhance promoter activity by −60%. Similarly, with the −85 CAT construct, BSIF-1 and BSIF-3 had a 2−3-fold negative effect in CaCo-2 cells but no significant effect in HepG2 cells. Finally, we tested the effects of BSIF-1 and BSIF-3 upon activity of a minimal promoter, extending from −30 to +121 (TATA-CAT). A 2-fold negative effect on this minimal promoter was seen when four copies of BSIF-1 were linked to it, both in CaCo-2 and HepG2 cells. BSIF-3 exhibited a weak positive effect in CaCo-2 cells and a weak negative effect in HepG2 cells (Fig. 2).

From the data with the promoter deletion constructs, we conclude that BSIF-1 strongly represses transcription of the strong −139 CAT apoB promoter in CaCo-2 cells (12-fold). Repression by this oligonucleotide appears to be independent of orientation, similar to an enhancer or silencer. There is clearly a target for this repression by BSIF-1 in the apoB promoter (somewhere between −139 and −1). Studies with both the −111 CAT promoter, the −85 promoter and with the −30 CAT promoter showed only a 2−3-fold repression by BSIF-1 in CaCo-2 cells, suggesting that the principal target for repression may not be localized between −111 and −30. The weak, 2-fold repression seen with the −85 CAT construct and the −30 CAT construct may be due to an interaction of BSIF-1 with sequences from −85 to −1 or to a direct interaction of BSIF-1 with components of the basal transcriptional machinery. The loss in repression (−5-fold) seen with the −111 CAT construct as compared to the −139 CAT construct suggests that a target for repression by BSIF-1 and BSIF-3 in CaCo-2 cells may reside between −139 and −111 of the apoB promoter.

To address these questions, we prepared a series of constructs designed to shed light upon the mechanism of repression by BSIF-1 and BSIF-3 in CaCo-2 cells. First, we asked whether the segment from −139 to −111 contains a target sequence for the repressor. To this end, three copies of an oligonucleotide harboring the sequence from −139 to −111 of the apoB promoter were cloned downstream of the TATA-CAT construct to generate the plasmid (−111−139)−TATA-CAT. Upstream of this construct were cloned three and four copies of BSIF-1 and BSIF-3, respectively (Fig. 3). Transient transfections of these constructs into CaCo-2 cells revealed that BSIF-1 repressed the activity of the (−111−139)−TATA-CAT construct by about 6-fold, similar to that observed with the −139 CAT construct. In contrast, BSIF-3 did not repress, but rather enhanced, activity of the (−111−139)−TATA-CAT construct, suggesting that the mechanism by which BSIF-1 and BSIF-3 influence apoB promoter activity differs. In CaCo-2 cells, the level of repression seen with the −111 CAT construct was similar to that observed with the −85 CAT construct (−2-fold) (Fig. 2), suggesting that sequences between −111 and −85 of the apoB promoter may not be involved in the repression effect. That this is indeed the case is demonstrated in Fig. 3. Three copies of an oligonucleotide representing the sequence from −111/−85 were cloned upstream of TATA-CAT to create the (−111−85) TATA-CAT construct. BSIF-1 did not repress this promoter, but rather slightly enhanced its activity, as was also the case with BSIF-3; thus establishing that the segment of the apoB promoter from −111 to −85 is not implicated in the repression by BSIF-1.

From the data in Fig. 3, it would appear that most of the repressor effect of BSIF-1 in CaCo-2 cells is attributable to an interaction with the segment from −139 to −111 of the apoB promoter, with a secondary and minor target present between −85 and the transcriptional start site (Fig. 2). A C/EBP-binding site is localized in the sequence from −53 to −72 of the apoB promoter. To eliminate it as a possible target of repres-
transfection of these constructs into CaCo-2 cells (Fig. 3) revealed no repression by BSIF-1 and a very minor repression by BSIF-3 in CaCo-2 cells, thus demonstrating that the C/EBP-binding site situated between -72 and -53 of the apoB promoter plays no role in the repression effect of BSIF-1 in CaCo-2 cells. This result, together with the data in Fig. 2, showing that the repression effect of BSIF-1 upon the -85 CAT and the TATA-CAT constructs is about the same (2-fold), suggests that the minor repression effect seen with these two constructs results from an interaction with the basal transcriptional machinery. Therefore, we conclude that BSIF-1 repression of apoB promoter activity in CaCo-2 cells involves two targets, a principal target in the sequence from -139 to -111 and a secondary target in the vicinity of the transcription initiation complex.

On the other hand, repression of the apoB promoter by BSIF-3 in CaCo-2 cells does not appear to involve a single main target but rather to require the full integrity of the -139 CAT promoter for maximal effect. As shown in Fig. 2, in CaCo-2 cells, the -139 CAT construct is repressed 10-fold by BSIF-3, while the -111 CAT construct is repressed only 2-3-fold, as also seen for BSIF-1. Similarly, the level of repression by BSIF-3 remains 2-3-fold upon the -85 CAT construct. In contrast with the results seen with BSIF-1, BSIF-3 exhibits no repressor activity at all when situated upstream of the -139/-111 TATA-CAT construct; thus eliminating this segment (-139 to -111) as a putative BSIF-3 target (Fig. 3). Similarly, the sequence from -111 to -85 is not repressed by BSIF-3. A minor (-30%) repressor effect is imposed by BSIF-3 upon the C/EBP-TATA-CAT construct but no effect of BSIF-3 at all is seen upon TATA-CAT. Therefore, the combined data concerning BSIF-3 in CaCo-2 cells suggests that the entire segment between -139 and -30 of the apoB promoter is required for BSIF-3 to exert its repressive effect. In CaCo-2 Cells, the Repressor Effect of BSIF-1 and BSIF-3 Cannot Be Overcome by the ApoB Second Intron Enhancer—The following experiment was designed to gain insight into the mechanism of repression by BSIF-1 and BSIF-3. Because the magnitude of the negative effect is maximal upon the -139 CAT promoter construct (Fig. 2), we used this construct to ask whether the strong repression caused by BSIF-1 and BSIF-3 is overcome by the presence of the apoB second intron enhancer. The apoB second intron enhancer has been extensively characterized in our laboratory and encompasses a 443-bp fragment (621–1064), harboring functional binding sites for the transcription factors HNF-1, C/EBP, and protein II (13).

In previous experiments of this nature involving a different repressor of apoB transcription, ARP-1, that binds to the 5’ upstream region of the apoB gene, we gathered evidence suggesting that the apoB second intron enhancer may function as do other classical enhancers, i.e. by interacting directly with the basal transcriptional machinery (24). Thus, we expected that if the repression effect promoted by BSIF-1- and BSIF-3-binding proteins also involved an interaction with the basal transcription complex, competition between the repressor and enhancer for available sites of interaction within the basal transcription complex may ensue. If enhancer and repressor proteins compete for interaction with the transcriptional machinery, and there is preferential occupancy of the machinery by the repressor, the enhancer might not be able to overcome this repression effect. Alternatively, if the mechanism of activity of the repressor does not involve the basal transcription complex, the enhancer should relieve the repression.

The 443-bp apoB enhancer was cloned upstream of the -139 CAT construct harboring BSIF-1- and BSIF-3-binding sites. As shown in Fig. 2, BSIF-1 and BSIF-3 decreased the CAT activity of the -139 CAT construct 10–12-fold in CaCo-2 cells. Addition of the enhancer increased the promoter activity of the constructs repressed by BSIF-1 and BSIF-3 only by 2-fold (Fig. 4), thus demonstrating that the enhancer is unable to overcome repression by BSIF-1 and BSIF-3. Based on our hypothesis, this implies that BSIF-1- and BSIF-3-binding proteins must exert their repression by means of a direct or indirect interaction with the basal transcription complex. Our data further suggests that the interaction between BSIF-1- and BSIF-3-binding proteins, their targets, and the basal transcription complex take preference over the interaction of the enhancer with the basal transcription complex.

Binding of BSIF-1 to Nuclear Proteins—To identify nuclear
proteins binding to BSIF-1, we performed gel retardation experiments. In Fig. 5, Panel A, lane 1, we observe that the BSIF-1 oligonucleotide forms two retarded complexes (A and B) with nuclear proteins from CaCo-2 cells, a result reminiscent of that of Traber et al. (14) with SIF-1. These investigators have shown that CaCo-2 proteins bind to the SIF-1 element either as a monomer (complex A) or a dimer (complex B) (see Fig. 5, Panel A, lane 5), and that dimer formation depends on the redox potential of the reaction, with homodimer formation being favored under reducing conditions (15). Formation of both the BSIF-1- and SIF-1-retarded complexes, A and B, is abolished in the presence of a 200-fold molar excess of unlabeled homologous oligonucleotides (lanes 3 and 7), but not by heterologous sequence (lanes 4 and 8, respectively). Furthermore, an antibody raised against mouse SIF-1 supershifts the retarded complex B formed with labeled SIF-1 (lane 6) but not with BSIF-1 (lane 2). The antibody appears to preferentially recognize the dimer. Unlike SIF-1 (Panel B, lanes 4–6), BSIF-1 can form retarded complexes with nuclear proteins from HepG2 cells (Panel B, lanes 1–3). These complexes are specific as shown in Panel B, lane 2, because of competition by an excess of unlabeled BSIF-1 oligonucleotide. Again, SIF-1 does not compete for the formation of complexes A and B in HepG2 cells (Panel B, lane 3).

Competition experiments using excess unlabeled oligonucleotides, as well as experiments using antibodies against the mouse SIF-1 protein, suggested that BSIF-1 and SIF-1 bind similar proteins with widely different affinities. That this is indeed the case is illustrated in Fig. 5, Panel C. Using CaCo-2 extracts, we observed that a 100-fold excess of homologous oligonucleotide competes for over 90% of the specific complexes formed with the SIF-1 probe (lanes 2 and 3). A 300-fold excess of homologous SIF-1 oligonucleotide nearly abolishes SIF-1 complex formation (lane 4), while a 300- and 500-fold excess of BSIF-1 oligonucleotide reduces SIF-1 complex formation by only 50 and 60%, respectively (lanes 5 and 6). Furthermore, a 500-fold excess of SIF-1 oligonucleotide eliminates BSIF-1 complex formation (lanes 7 and 8).

Recently, a cDNA for the mouse SIF-1-binding protein was cloned and sequenced. Analysis of the derived protein sequence revealed that the SIF-1-binding protein is a member of the caudal family of homeodomain proteins and was designated mCdx-2 (15). We then asked whether proteins in an extract made from COS cells overexpressing mCdx-2 would form retarded complexes with BSIF-1. For comparison, we used a CaCo-2 extract in the same experiment. As seen in Fig. 6, Panel A, both the BSIF-1 oligonucleotide (lane 1) and the SIF-1 oligonucleotide (lane 6) formed complexes A and B with the CaCo-2 cell protein. A control COS cell extract, transfected with the empty plasmid, formed no retarded complexes with either of the two probes (lanes 5 and 10). However, the extract from mCdx-2 expressing COS cells formed strong complexes with SIF-1 as expected (lane 7) that were abolished in the presence of an excess of unlabeled SIF-1 oligonucleotide (lane 8). Furthermore, these complexes were supershifted by the antibody against mCdx-2 (lane 9), as predicted. The interaction of mCdx-2 protein with the BSIF-1 oligonucleotide was much weaker (lane 2), although specific (lane 3), suggesting that the affinity of mCdx-2 for the BSIF-1 sequence is much lower than that for the sucrose isomaltase sequence. This result is in agreement with data in Fig. 5, Panel C. In summary, mCdx-2 does bind to BSIF-1, but with a much lower affinity as compared to its binding to SIF-1, suggesting that the major protein from CaCo-2 extracts that binds BSIF-1 may not be the human homolog of mCdx-2.

The cDNA for another member of the caudal family has been cloned, namely that for mCdx-4 (20). An mCdx-4 cDNA clone was kindly provided to us by Drs. Gamer and Wright. This allowed us to ask whether the BSIF-1 sequence binds to mCdx-4. As shown in Fig. 6, Panel B, lanes 3 and 4, a COS cell extract transfected with a mCdx-4 expression vector does form specific retarded complexes with BSIF-1. Formation of complex A is competed for by BSIF-1, while competition for complex B is incomplete. As a control, a sample of the cell extract trans-
fected with an empty plasmid did not yield the retarded complexes (lane 5). The retarded complexes observed with mCdx-4 are very similar to those seen with the CaCo-2 extracts (lanes 1 and 2). The oligonucleotide/nuclear extract binding reaction was incubated with an antibody against mCdx-4 (kindly provided by Drs. Gamer and Wright). Although no supershifted band was seen in the presence of the antibody, formation of the retarded complexes was diminished, particularly that of the dimer, suggesting that the antibody indeed recognizes the mCdx-4 protein and interacts with it in a manner that interferes with DNA binding (data not shown). Therefore, we conclude that the BSIF-1 sequence interacts with protein(s) of the caudal family, most likely with the human homolog to mCdx-4, and that binding of this protein to the BSIF-1 sequence can repress transcription from the apoB promoter.

Analysis of Proteins Binding to the Sequence of the ApoB Promoter between 2139 and 2111—Having determined that the repressor effect of BSIF-1 is due, at least in part, to binding by a human protein related to mCdx-4 (and to a lesser extent, mCdx-2), the next step was to identify any possible targets for the repressor. Since we have shown in Figs. 2 and 3 that the repressor effect is centered upon the segment from 2139 to 2111 of the apoB promoter, we examined whether nuclear proteins from CaCo-2 cells may bind to this region. To this end, an oligonucleotide representing the 2139 to 2111 sequence (139/111) was incubated with CaCo-2 nuclear extract and complex formation examined in gel retention experiments. One such experiment is shown in Fig. 7. In Panel B, lane 1 (left side), we observed two specific retarded complexes whose formation was competed by an excess of unlabeled probe (data not shown). At this stage, we wished to determine whether BSIF-1- and/or BSIF-3-binding proteins can also bind to the 139/111 sequence thereby exerting their repressive effect by simply competing with one or more activator(s) binding to that sequence. As shown in Panel B, lanes 2 and 3, neither BSIF-1 nor BSIF-3 oligonucleotides at a 200-fold excess were able to compete for complex formation between the CaCo-2 proteins and the 2139 to 2111 sequence, thus eliminating the possibility that repression of the 2139 to 2111 sequence by BSIF-1 and BSIF-3 may be due to competition for DNA binding. Mutagenesis of 6 bases within the 139/111 sequence obliterated DNA binding (Panel B, right side, lanes 1–3) and the repressor effect of BSIF-1 (Panel C), thus demonstrating that the segment between 2139 and 2111 is indeed a target for the repression by BSIF-1.

To identify the protein(s) binding to the 2139 to 2111 sequence, we first compared the sequence of this segment to the consensus binding site for the most common tissue-specific transcription factors, namely HNF-1, HNF-3, HNF-4, and C/EBP. A resemblance was found to the C/EBP-binding site. Kardassis et al. (25) had described that a binding site for C/EBP may be present in the region between 2118 and 2110 of the apoB promoter. Because the repressor effect of BSIF-1 upon the 2139 to 2111 sequence was stronger in CaCo-2 than in HepG2 cells, we reasoned that the 2139 to 2111 sequence may bind the intestine-enriched isoforms of C/EBP, namely -β and
This hypothesis was tested directly using C/EBP-β and -δ proteins derived from COS cell extracts transfected with the corresponding cDNA expression clones. As shown in Fig. 8, Panel B, the −139 to −111 oligonucleotide forms a major specific retarded complex with CaCo-2 cell extracts (lanes 1 and 2). The same complex is seen when extracts enriched in C/EBP-β (lanes 4 and 5) or -δ (lanes 6 and 7) are used, thus demonstrating that the segment between −139 and −111 of the apoB promoter does bind C/EBP-β and/or -δ. Experiments in Fig. 7 showed that mutagenesis of this sequence abolishes protein binding (Panel B, right side) and repressor activity (Panel C), thus strongly suggesting that the repressor effect of BSIF-1 involves an interaction with members of the C/EBP family binding to the −139 to −111 region of the apoB promoter.
C/EBP-\(b\) and -d indeed activate this segment of the promoter is shown in Fig. 8, Panel C. Cotransfection of a -139 to -111/TATA-CAT plasmid with either C/EBP-\(b\) or -d did increase transcriptional activity in CaCo-2 cells by a factor of two, consistent with our hypothesis. The increase seen is not greater because CaCo-2 cells are not depleted in these C/EBP proteins. These results are significant because they show for the first time that caudal proteins, known so far as activators of transcription, can also repress transcription and that at least one component of the repressor effect involves an interaction with C/EBP proteins. The intricate details of the molecular interactions between the BSIF-1 sequence, its protein, and the C/EBP
binding region between $-139$ to $-111$ remain the object of future studies.

**BSIF-3 and SIF-3 Bind HNF-1 with Different Affinities**—To identify proteins binding to BSIF-3, we performed gel retardation experiments. In Fig. 9, Panel A, we show that BSIF-3 forms a similar retarded complex with nuclear proteins from mouse liver (lane 1) and CaCo-2 cells (lane 2). The retarded complex is competed for by a 200-fold molar excess of BSIF-3 oligonucleotide, as well as by the SIF-3 oligonucleotide (lanes 3 and 4), respectively, suggesting that regulatory sequences from the apoB gene and the sucrose isomaltase gene may be bound by the same protein. That this indeed is the case is demonstrated in lanes 5–8; SIF-3 forms a retarded complex with nuclear extracts from CaCo-2 cells and mouse liver of the same mobility as the complex found by BSIF-3 (lanes 1 and 2). Formation of the SIF-3 complex is abolished by a 200-fold excess of unlabeled homologous oligonucleotide (lane 6). However, a similar excess of unlabeled BSIF-3 oligonucleotide competes only partially for binding of SIF-3 to its protein (lane 7). Yet we showed that SIF-3 competes well for binding of BSIF-3 to its protein. These two results can be explained by a different affinity for binding of the same protein to BSIF-3 and SIF-3 DNA, with affinity of the protein being much higher for the SIF-3 oligonucleotide than for BSIF-3. That is indeed the case is demonstrated by the results in Fig. 9, Panel B. The SIF-3-retarded complex formed with CaCo-2 nuclear proteins (lane 1) is readily abolished by the presence of a small excess (25-fold) of unlabeled homologous oligonucleotide (lane 2). In contrast, the BSIF-3 oligonucleotide competes poorly for SIF-3 complex formation (lanes 4–6), even at a 500-fold excess. On the other hand, BSIF-3 complex formation with either CaCo-2 cells or mouse liver proteins is readily competed for by only a 100-fold excess of SIF-3.

Recently, Wu et al. (26) have shown that SIF-3 binds to HNF-1α. We have confirmed this result and have also shown that BSIF-3 binds to HNF-1α by using an extract from COS cells expressing HNF-1α (Fig. 10, lane 2). The BSIF-3-HNF-1α complex is supershifted by an HNF-1 antibody (lane 4), further confirming that BSIF-3, like SIF-3, binds HNF-1.

**DISCUSSION**

To date there is little information regarding the DNA sequences and transcription factors that participate in the transcriptional activation of the apoB gene in intestinal cells in vivo, although much is known concerning transcription in hepatic cells (for review, see Ref. 4). Studies of the distribution of apoB mRNA along the intestinal tract show that an apoB mRNA abundance is similar in duodenum, jejunum, and ileum. In every region of the small intestine the mRNA concentration is lower in crypt compared to villus epithelium (27). Developmental studies of apoB gene expression in liver and intestine performed using mRNAs derived from fetal, suckling (postnatal days 1–12), weaning (postnatal days 13–28), and adult (250 g, 70-day-old) rats demonstrated that apoB mRNA is as abundant in 18-day fetal liver as in any subsequent period of development. In contrast, the concentration of apoB mRNA remains low in fetal intestine until the last (21st) day of gestation, when it increases sharply to levels that are severalfold higher than in the liver (27). The postnatal pattern of apoB mRNA accumulation is similar in liver and intestine. The low levels of intestinal apoB mRNA prior to the last 2 days of gestation may correlate with the pattern of gut cellular differentiation. Significant changes in apoB mRNA accumulation have been observed during the postnatal period. For example, an increase in small intestinal apoB mRNA concentrations during the first postnatal week parallels the high rates of lipid absorption known to

---

**Fig. 9.** Binding of the BSIF-3 and SIF-3 oligonucleotides to nuclear proteins from CaCo-2 cells and mouse liver. Panel A, on top of the gel are shown the probe, the extract, and the competitor oligonucleotide used in each reaction. The retarded complexes are also indicated. In Panel B we show the relative binding affinities of BSIF-3 and SIF-3 oligonucleotides to nuclear proteins.
ApoB Gene Repression in Intestinal Cells

---

Fig. 10. HNF-1α binds to the BSIF-3 oligonucleotide. The supershifted complex is designated S.

---

... occur in the suckling rat (28). Furthermore, there is a pronounced decline in liver and intestinal apoB mRNA levels during weaning and juvenile development. Factors responsible for these large decreases in hepatic and intestinal apoB mRNA concentration and for their subsequent increase during later adulthood remain to be investigated in detail.

The epithelium of the intestinal tract arises from the visceral endoderm by progression through a series of developmental transitions (for review, see Refs. 29 and 30). Early influences upon intestinal development are partially directed by the interaction of mesoderm with endoderm, although the molecular mechanism of this effect has not been elucidated (31). Later events include the emergence of different epithelial cell lineages, formation of crypts (proliferative compartment) and villi (differentiated compartment), and subsequent coordinated expression of cell-specific genes. Mature epithelia of the small intestine and colon reveal marked regional differences in architecture, cell populations, and gene expression defining distinct functional zones within the adult gut (32–34). Such region-specific phenotypes persist throughout the life of the animal. Experiments with transgenic mice have proved that developmental signals for gene expression in the intestinal epithelium derive from regulatory elements of several intestinal genes (32, 35, 36). Analysis of the regulatory regions of the intestinal and liver fatty acid-binding proteins and sucrose-isomaltase genes reveal poorly defined positive and negative elements that direct cell lineage and spatial patterns of expression. Analysis of these three genes implies that the process by which expression is restricted to the intestine is regulated at the level of transcription (32, 35, 36).

CaCo-2 cells have been used as a model system to study gene expression in intestinal cells. These cells display many functions characteristic of fully differentiated small intestinal mucosal epithelium (37). They express brush border hydrolases (38), gut-neuropeptide hormone responsiveness (39), vitamin uptake (40), and ion transport (41) and exhibit structural characteristics typical of the small intestinal enterocyte, such as brush border microvilli, tight junctions, and dome formation on an impermeable support (42). CaCo-2 cells synthesize and secrete all the components of the major classes of human plasma lipoproteins and thus represent a useful in vitro system for the study of the regulation of lipoproteins and apolipoprotein synthesis. Transient transfection experiments of apoB promoter segments into CaCo-2 cells have shown that only 139 bp of upstream sequence are sufficient for high level promoter activity in these cells (11).

Recent work with the sucrose isomaltase gene that is expressed exclusively in intestinal cells has defined a new transcription factor involved in activation of this gene in CaCo-2 cells, namely mCdx-2, a member of the caudal family of homeodomain proteins. mCdx-2 binds to the segment from −54 to −32 of the SI promoter to activate transcription (15). mCdx-2 is a member of a family of homeodomain proteins related to the caudal (cad) gene of Drosophila melanogaster (43). Genes related to caudal have been found in many other species such as mice (44–46), Syrian hamsters (47), rats (48), chickens (49), Xenopus laevis (50), zebra fish (51), and Caenorhabditis elegans (52). Genes related to caudal have functional interactions with regulatory elements of the insulin (47) and fushi tarazu (ftz) genes (53). The fact that the Cdx-2 protein plays an important role in transcription activation of the SI gene raises the issue of whether Cdx proteins play a more general role in the regulation of other intestinal genes such as apoB and in directing intestinal development.

Indirect evidence exists for a role for Cdx genes in vertebrate intestinal development. In mice, mCdx-1 mRNA and protein are first expressed in the ectoderm and mesoderm of the primitive streak and in derivatives of these tissues until approximately embryonic day 12, after which expression appears to be extinguished (54). Before the endoderm-intestinal epithelial transition (day 14), mCdx-1 is expressed in the endoderm and later in intestinal epithelium (54). Mouse Cdx-4 mRNA and protein are expressed early in gastrulation but not later in development (20). The bimodal pattern of expression noted for genes related to caudal during development is consistent with the developmental expression of other homeobox genes. The patterns of Cdx gene expression in endoderm and intestine suggest a developmental role for this gene.

In this report, we describe a novel function for caudal proteins, namely that of repressors of transcription of the apoB gene in intestinal cells. Based on our observation that within intron 3 of the apoB gene there is a sequence highly homologous to the binding site for mCdx-2 in the SI gene that is responsible for SI gene activation, we focused on this region and asked whether mCdx-2 interacts with the apoB sequence. Our results showed that both mCdx-2 and mCdx-4 (another member of the caudal family of proteins), bind to the apoB sequence, with mCdx-4 binding showing a higher affinity than mCdx-2. The interaction of caudal homeodomains with DNA sequence elements was examined in detail for the chicken protein CdxA (55). Using both random oligonucleotides and genomic DNA fragments, DNA sequences that bind to CdxA were selected. In this manner, a consensus sequence for Cdx binding emerged, namely (A/C)TTTAT(A/G). Work by Suh et al. (15) with the sucrose isomaltase gene showed that the SIF-1 element is palindromic in nature, consisting of two half-sites, each of 7 bases in length. These two half-sites, namely TTTATTT and CTTTATG, correspond very well to the consensus binding sequence for caudal homeodomain proteins. Studies of the binding of mCdx-2 to the SIF-1 sequence demonstrated that a protein of Mr = 38,000 binds preferentially to the
5' half-site of the SIF-1 element. Occupation of the second half-site of the SIF-1 element leads to a higher molecular weight complex representing a homodimer of mCdx-2. Formation of the dimer of Cdx-2 on the SIF-1 element is dependent on the redox state of the binding reaction; with dimer formation favored under reducing conditions, in the presence of dithiothreitol (15). The mechanism of the redox effect on protein binding is unknown but the presence of two conserved cysteines situated upstream of the conserved homeodomain suggests that they may play a role in the protein-protein interactions that are affected by the redox state.

The putative caudal-binding site within the BSIF-1 element differs from that of the SIF-1 element. There are two nucleotide differences between the 5' half-site of SIF-1 and the corresponding sequence in BSIF-1 (AATAAAA versus AATAGG). However, in the 3' half-site, there is identity at only 4 bp out of 7 between SIF-1 and BSIF-1 (CTTTATG versus TccATG). The 2-bp difference between the 5' half-site of SIF-1 and BSIF-1 may account for the diminished binding affinity of BSIF-1 as observed in Fig. 5. A mutation in the 5' half-site of the SIF-1 element, AATAACC (M4 in Ref. 15), revealed that overall binding to intestinal-derived nuclear proteins was diminished and dimer formation abolished, suggesting that the integrity of the 5' half-site is important for binding at the 3' half-site. Additional experiments demonstrated that the affinity for binding to the 5' half-site of the SIF-1 element is much higher than the affinity for the 3' half-site. Indeed mutations in the 5' half-site of the SIF-1 element were just as detrimental to the transcriptional activity as were mutations in both the 5' and 3' half-sites of the SIF-1 element, demonstrating that binding to the 5' site of the SIF-1 element stimulates transcription of the SI gene. The importance of the 5' half-site in the binding and function of the SIF-1 element implies that the similar sequence within the BSIF-1 element may play a role in binding and in the repression effect. Unlike the situation with the SIF-1 element, mCdx-2 (or its human homolog) may not be the major binding protein responsible for the retarded complexes with BSIF-1 seen both with CaCo-2 and HepG2 nuclear extracts; indeed our results (Fig. 6, Panel B) suggest that mCdx-4 (or its human homolog) may be the major protein binding to the BSIF-1 element.

The protein sequences of mCdx-2 and mCdx-4 differ in two important aspects; first, all caudal-like genes are classified according to their divergent version of a short, conserved hexapeptide motif found upstream of the homeodomain. The mCdx-4 hexapeptide is very similar to those of other vertebrate caudal-like genes. However, only mouse mCdx-2 and hamster Cdx-3 diverge from this conserved caudal sequence, by having a cysteine residue instead of a tyrosine in the second position of the six-amino acid core. Second, unlike mCdx-2, that has conserved tyrosine and serine residues in the amino-terminal portion of the homeodomain, mCdx-4 contains unique cysteine and asparagine residues. These differences among the sequences of mCdx-2 and mCdx-4, near or at the homeodomain, may be responsible for the increased affinity of mCdx-4 for the BSIF-1 element as compared to mCdx-2. This hypothesis will be tested experimentally in the future. The BSIF-1 sequence caused a 12-fold repression of transcription of an apoB promoter construct (−139 CAT) in CaCo-2 cells but not in HepG2 cells. The target within the apoB promoter for this repressor effect appears to be the sequence from −139 to −111, where a C/EBPβ- and/or -β-binding site resides. The interaction between mCdx-4 (or -2) and the −139 to −111 region probably involves an alteration in the architecture of the basal transcription complex, as suggested by the results of experiments in which the apoB second-intron enhancer, that normally interacts with the basal transcription complex, is unable to neutralize the effect of the repressor (Fig. 4). The much weaker repression observed in HepG2 cells as compared to CaCo-2 cells may be due to differences in the architecture of the transcription complex in these two cells. Alternatively, it may be that the tissue-specific differences in C/EBP isoforms (that are the target for repression between these cells) may account for the observed differences in repression.

Our data show that members of the caudal family of proteins can function as transcriptional repressors as well as activators as described previously (14). The idea of a single DNA sequence element serving as a target for both activators or repressors is not novel. For example, both the liver-enriched transcription factor HNF-4 and ARP-1, a member of the family of steroid-thyroid nuclear receptors, bind to the same DNA sequence in the apoB promoter. Binding of HNF-4 leads to transcriptional activation while binding of ARP-1 leads to transcriptional repression (56). Another example is provided by ftz and engrailed (En), two of the genes required for proper segmentation of the Drosophila embryo. Their protein products each contain a homeodomain and recognize the same DNA sequences located upstream of the TATA box in the hsp70 promoter. Binding of ftz to this DNA sequence activates transcription while En represses this activation by competing with ftz for binding to these sites (57).

Unlike the sucrose isomaltase gene, where mCdx-2 binds only to the SIF-1 sequence in intestinal cells (14), proteins related to mCdx-4 and mCdx-2 present in HepG2 cells also bind to the BSIF-1 sequence, suggesting that these proteins may also play a role in the regulation of hepatic-specific genes. Questions remain as to whether the repressor effect exerted by BSIF-1 upon the −139 to −111 sequence reflects an active mechanism of repression involving mCdx proteins and C/EBPs or -β or -δ. Further, are there other regions in the apoB gene where the caudal family members may play a positive role in transcription? Future studies will attempt to answer these questions.

HNF-1 is also a member of a class of transcription factors containing a DNA binding domain that is related to that of homeobox proteins (5). An HNF-1-binding site within the intron 2 enhancer of the apoB gene functions as an activator of transcription in both HepG2 and CaCo-2 cells (17). Interestingly, binding of HNF-1 to the BSIF-3 sequence situated in intron 3, causes a strong decrease in apoB promoter activity in CaCo-2 cells and a weaker one in HepG2 cells. Although HNF-1α appears to be the major protein binding to BSIF-3, we cannot discard the possibility that another protein present at low concentration in our extracts may bind with high affinity and cause the repression observed. In contrast with the situation with BSIF-1, where a target for the repression has been identified in the segment from −139 to −111 of the apoB promoter, it would appear that the integrity of the region from −139 to −85 is necessary for the effect of BSIF-3 upon transcriptional activity.

Acknowledgments—We are grateful to Drs. Peter Traber and Eu-Ran Suh for providing us with mCdx-2 cDNA and antisera and for sharing with us results prior to publication, to Drs. Laura Gamer and Christopher Wright for providing us with mCdx-4 cDNA and antisera, and to Dr. Robert Costa for donating the HNF-1α expression plasmid and antibody. We give special thanks to Drs. Brian J. McCarthy and Allen D. Cooper for critically reading this manuscript, and to Rick Cuevas for manuscript preparation assistance.

REFERENCES
1. Xantopoulos, K. G., and Mirkovitch, J. (1993) Eur. J. Biochem. 216, 353-360
2. Galson, D. L., Tschiya, T., Tender, D. S., Huang, L. E., Ren, Y., Ogura, T., and Bunn, H. F. (1995) Mol. Cell. Biol. 15, 2135-2144
3. Rouet, P., Raguenes, G., Trousche, F., Mfoudou, V., and Salière, J. P. (1995) Nucleic Acids Res. 23, 395-404
