Identification of a Goblet Cell-specific Enhancer Element in the Rat Intestinal Trefoil Factor Gene Promoter Bound by a Goblet Cell Nuclear Protein

Haruhiko Ogata, Nagamu Inoue, and Daniel K. Podolsky†

From the Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

Intestinal trefoil factor (ITF) is selectively expressed in goblet cells of the small and large intestinal mucosa. Detailed analysis of the rat ITF (RITF) promoter was undertaken by transient transfection and gel mobility shift assays (GMSAs) using the goblet cell-like LS174T colon cancer-derived cell line. Various lengths of wildtype or mutant constructs of the 5′-flanking region were linked to the pXP2 reporter gene luciferase. Expression of −118 RITF was significantly decreased compared with −154 RITF, and transfection with an 18-base pair construct (−141 to −124) resulted in more than 5-fold greater expression than transfection with the promoterless pXP2 gene construct alone. Using various synthetic oligonucleotide mutants, GMSAs revealed that only a 9-base pair sequence (CCCCCTCCCC) in this element was required for specific binding, overlapping but distinct from a Sp1-like element. GMSA demonstrated that this element was specifically bound by nuclear proteins from intestinal cells with a goblet cell-like phenotype. These studies demonstrate that a 9-base pair element (goblet cell response element) between −154 and −118 in the RITF promoter gene is a cis-active element bound by a distinct nuclear transcription factor and is capable of directing intestine and goblet cell-specific expression.

Trefoil factors are a family of small peptides expressed at various sites throughout the gastrointestinal tract. The members of this family share an array of structural features including a distinctive motif of six cysteine residues termed a trefoil or a P domain. Thim (1) postulated that the six cysteine residues could contribute to the formation of three intrachain loops via the formation of disulfide bonds; the resultant predicted three-looped structure prompted the trefoil designation. A recent nuclear magnetic resonance analysis as well as x-ray crystallography of one of the trefoil peptides supported the presence of a distinctive secondary structure consistent with the putative three-intrachain loop formation (2). Members of the family identified in mammals possess one or two P domains (3–5). Amphibians have been found to express trefoil proteins with as many as four P domains (6).

Members of the trefoil peptide family appear to be expressed in a region-specific fashion along the length of the gastrointestinal tract. Human spasmolytic polypeptide bears two trefoil motifs and is expressed primarily in the stomach (5), although the porcine homologue was originally isolated from pancreas (7), pS2, bearing a single trefoil motif and initially cloned as the product of an estrogen-responsive gene from a breast cancer cell line (8), is normally expressed only in the gastric antrum in man (9). Cloning of the rodent homologues of pS2 and spasmolytic polypeptide confirmed that expression is site-specific along the longitudinal axis of the upper gastrointestinal tract in a pattern that has been conserved in evolution (10, 11).

Intestinal trefoil factor (ITF) is a third member of the trefoil peptide family identified in humans (12, 13), rats (14, 15), and mice (16, 17) that contains a single P domain. In contrast to pS2 and human spasmolytic polypeptide, ITF is normally selectively expressed in the normal small and large intestinal mucosa, complementing the pattern of expression of the other members of the family in the normal gastrointestinal tract. More specifically, ITF expression is normally confined to the goblet cell population within the intestinal epithelium (14).

Goblet cells of the small and large intestine secrete a complex mixture of mucin glycoprotein onto the cell surface, but their functional importance in gastrointestinal tract mucosa has not been well defined. Moreover the basis of selected gene expression responsible for the distinctive goblet cell phenotype has not been defined. Genes encoding the apomucin peptide backbones of mucin glycoproteins are enormous in size and highly complex, hampering progress in efforts to define the regulatory effect conferring goblet cell-specific expression. The selective expression of ITF in intestinal goblet cells suggests that characterization of the gene encoding this peptide may provide insight into regulatory elements responsible for goblet cell-specific gene expression. Among the trefoil family members, only the promoter of pS2 gene normally expressed in gastric mucosa has been partially characterized (18, 19). Although several genes expressed in the intestinal epithelium including fatty acid-binding protein (20), sucrase-isomaltase (21), and lactate (22) have been cloned and their regulatory elements studied, none are products of goblet cells.

Our previous report showed relatively high levels of specific expression in transient transfection studies using promoters as short as 154 bp of the 5′-flanking region of the rat ITF (RITF) gene. The presence of goblet cell-specific promoter element(s) within close proximity to the transcriptional start site was suggested by those preliminary efforts (15). Further investigation of the RITF promoter was undertaken, and a cis-active element capable of directing goblet cell-specific expression was identified in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Gastrointestinal Unit, Massachusetts General Hospital, Fruit St., Boston, MA 02114. Fax: 617-724-2136; E-mail: Podolsky.Daniel@mgh.harvard.edu.

‡ The abbreviations used are: ITF, intestinal trefoil factor; RITF, rat intestinal trefoil factor; GCRE, goblet cell-response element; GMSA, gel mobility shift assay; EGF, epidermal growth factor; bp, base pair(s); WT, wild type; RSV, Rous sarcoma virus.

* These studies were supported by National Institutes of Health Grants DK46906 and DK43351. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

3060 This paper is available on line at http://www.jbc.org
identified within this region. Moreover, this element appears to be bound by a distinct nuclear transcription factor.

MATERIALS AND METHODS

Plasmid DNA Constructs—The promoterless luciferase gene construct pXP2 (23) was a gift from Dr. Lee Kaplan. The 5’-flanking region of the RITF gene, the 1671-bp fragment, was subcloned from pRITF20C containing the entire RITF gene into pXP2 vector to form the construct −1671 RITF-luc (1.7WT-luc) and transformed into competent E. coli DH5α cells (CLONTECH Laboratories; Palo Alto, CA) as described previously (15). To confirm correct orientation and preservation of the start codon of luciferase, plasmid DNA was subjected to restriction mapping and sequencing of the insertion junctions.

Deletion constructs of the 5’-flanking region of the RITF gene driving the luciferase gene, −979 RITF-luc (1.0WT-luc) were derived from the −1671 RITF-luc construct, taking advantage of convenient restriction sites as described previously (15). Further deletion plasmids, −664 RITF-luc (0.7WT-luc), −336 RITF-luc (0.3WT-luc), −154 RITF-luc (0.2WT-luc), and −117 RITF-luc (0.1WT-luc), were constructed by treatment of −1671 RITF-Luc with exonuclease III and mung bean nuclease after creation of a 5’-overhang using BamHI digestion to cut the plasmid at the 5’-end of the promoter insert using commercially available reagents (Erase-a-Base System, Promega, Madison, WI). −77 RITF-luc (0.05WT-luc) and −46 RITF-luc (0.02WT-luc) were generated by ligation of KpnI and BglII-digested polymerase chain reaction product derived from the 1.7WT-luc into the restriction site of the pXP2 vector. Minimal promoter DNA constructs containing wild-type (WT) or mutant elements (M) spanning −141 to −118 (WT1-luc) and −141 to −124 (WT2-luc or M3 to M8-luc) were generated by ligation of kinased double-stranded synthetic oligonucleotides into the BamHI site of the pXP2 vector.

The other deletions and mutants using full-length RITF promoter (1.7WT-luc) were prepared by replacing the wild-type sequence with deleted or mutated sequences generated by Transformer site-directed mutagenesis kit (CLONTECH). 1.7A1-luc represented deletion mutant with the deleted sequences from −154 to −118 and 1.7A2-luc from −141 to −124. The nucleotides for mutagenesis were chosen from the sequence between −141 and −124, and those mutated constructs were designated as 1.7M3- to 1.7M8-luc.

All constructs were verified by DNA sequencing using the Sequenase version 2.0 DNA sequencing kit (Amersham Life Science, Inc.). Plasmid preparation purity was confirmed by A260/A280 of > 1.6, and supercoiling of DNA was established by the appearance of agarose gel electrophoresis prior to use in transfection experiments.

Cell Cultures—Human colon cancer cell line LS174T and Caco-2, rat intestinal epithelial cell line IEC-6, human hepatocellular carcinoma cell line HepG2, and human cervix epithelial cell line HeLa obtained from the American Type Culture Collection (ATCC, Rockville, MD) were grown in Eagle’s minimum essential medium except IEC-6 in Dulbecco’s modified Eagle’s medium supplemented with 4.5 g/liter D-glucose, 10% fetal calf serum, Eagle’s balanced salt solution, nonessential amino acids, sodium pyruvate, 4 mM l-glutamine, 50 units/ml penicillin, and 50 μg/ml streptomycin. The H2 subclone of human colon cancer cell line HT-29 originally obtained from Dr. Daniel Louvard was grown in either Dulbecco’s modified Eagle’s medium with glucose (un-differentiating conditions) or in the presence of galactose as the sole source of carbohydrate to induce goblet cell-like differentiation as described previously (24, 25). All cell lines were grown in 5% CO2 at 37 °C.

Transient Transfection Promoter Analysis—Transient transfection was accomplished by the calcium phosphate precipitation method. Sixteen hours prior to transfection, 8 × 10^5 cells were plated out in triplicate in 35-mm wells of a six-well cell culture plate. Complete media was refreshed 2 h prior to transfection. Efficiency of transfection was standardized by co-precipitation of the construct of interest with pTK-GH, consisting of the minimal thymidine kinase promoter driving the human growth hormone gene as a reporter gene (26), and adjusting for the amount of human growth hormone expressed, as determined by a commercially available radioimmunoassay (hGH Allegro Kit, Nichols Institute Diagnostics, San Juan Capistrano, CA). Calcium phosphate-precipitated plasmid DNA was added to each well and incubated at constant 5% CO2 for 4 h before a 2-min exposure to 15% glycerol. Cells were subsequently cultured for 48 h prior to assay for reporter gene expression. For determination of luciferase activity, cells were lysed and assayed immediately using a commercial luciferase assay system (Promega) measured in a luminometer (Analytical Luminescence Laboratory, Monolite 2010). Luciferase activity was adjusted for transfection efficiency reflected in the level of growth hormone, expressed as nanograms of hGH/ml of medium. Where noted, promoter activity was expressed as a percentage of the expression of the maximal promoter construct RSV-luc (a gift from Dr. Loyal Tillotson), consisting of the RSV promoter joined to the luciferase gene, or a -fold increase of the expression of the pXP2.

Nuclear Extracts and Gel Mobility Shift Assays (GMSAs)—Nuclear extracts from cultured intestinal cells were prepared by Nonidet P-40 detergent lysis and 0.5 M NaCl extraction as described by Schreiber et al. (27). Nuclear extracts from other cell lines for GMSA tissue distribution studies were a kind gift from Drs. Anil K. Rustgi and Timothy C. Wang. Protein concentration was determined according to the Bradford
assay (28). The wild-type, double-stranded synthetic probes used in this study were \(^{15}T^{15}T^{15}T^{15}T^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C^{15}C
that a 9-bp palindromic sequence (band in GMSA that is not affected by Sp1 competitors suggests bind to the Sp1 binding site. Taken together, the presence of a other member of the Sp family did not affect either band. These by the antibody was not observed. Antibody against Sp3, an-
Sp1 could imperfectly abolish the upper band, and supershift B
). Antibody against Sp1 did not affect either complex (Fig. 4

complexes (Fig. 4B). The addition of Sp1-binding oligonucleotides to the binding mixture competed with formation of one protein complex (upper band), but not the other major complex formed between nuclear proteins and GCRE. The addition of mutated Sp1 did not affect either complex (Fig. 4B). Antibody against Sp1 could imperfectly abolish the upper band, and supershift by the antibody was not observed. Antibody against Sp3, another member of the Sp family did not affect either band. These findings suggest that the upper band may reflect the presence of a member of the Sp family (Sp1-like, not Sp3) that is able to bind to the Sp1 binding site. Taken together, the presence of a band in GMSA that is not affected by Sp1 competitors suggests that a 9-bp palindromic sequence (\^{135}CCCCTCCCC\textsuperscript{127}) of GCRE that overlaps but is distinct from an Sp1-like element in the RITF gene promoter is specifically bound by a goblet nuclear factor.

Gene Expression Analysis of the Mutated RITF Gene Promoter—To determine whether binding of GCRE in GMSA correlates with enhancer activity of this element, the effects of various mutations in \(-154\) to \(-118\) nucleotides of the RITF promoter gene expression were assessed by transient transfection assay. Minimal DNA constructs containing the wild-type 37-bp element (WT1, \(-154\) to \(-118\)) or 18-bp element (WT2, \(-141\) to \(-124\)) and the mutant 18-bp element (M3 to M8, \(-141\) to \(-124\)) were generated using synthetic oligonucleotides sub-
cloned into pXP2, and the resulting constructs were designated WT1-luc, WT2-luc, or M3- to M8-luc, respectively (Fig. 5A). Constructs were transiently transfected into LS174T cells, and gene expression was analyzed by measuring relative luciferase activity and represented as \(-fold increase of the activity of pXP2 (Fig. 5A). Transfections with WT1-luc and WT2-luc resulted in more than 5-fold greater (5.8-fold with WT1-luc and 5.2-fold with WT2-luc) expression than transfection with pXP2 alone. There was no significant difference in expression between WT1-luc and WT2-luc (Fig. 5A). Although mutations M3, M4, and M8 had no effect on wild-type WT2-luc expression, mutant constructs (M5-, M6-, and M7-luc) reduced expression nearly to the base line observed with the minimal construct pXP2 (Fig. 5A). The mutated sequences of M5 to M7 corresponded to the same 9-bp (\(-135\) to \(-127\)) element of GCRE identified through GMSA (Fig. 4).

The role of the GCRE in the context of the full-length promoter was assessed through the preparation of full-length constructs containing mutations in the region \(-154\) to \(-118\) analogous to the mutants prepared in association with the minimal promoter. Constructs containing 3’-deletion mutants with the deleted sequences (1.7\Delta1-luc deleted \(-154\) to \(-118\), and 1.7\Δ2-luc deleted \(-141\) to \(-124\) and those (1.7M3- to 1.7M8-luc) mutated in the same positions as represented in M3- to M8-luc indicated in Fig. 5A were generated by site-directed mutagenesis assay as described under “Materials and Methods” (Fig. 5B). After transfection into LS174T cells, promoter activities were analyzed and expressed as a percentage of the expression of the maximal promoter construct RSV-luc (Fig. 5B). As expected, the expression of the deleted mutants (1.7\Δ1-luc and 1.7\Δ2-luc) and the expression of the mutants M5 to M7 were
significantly decreased compared with the wild-type construct (1.7WT-luc) or constructs mutated at other sites within the 18-bp element (1.7M3-, 1.7M4-, and 1.7M8-luc) (Fig. 5B). These results were compatible with results from studies using the minimal DNA constructs and further implicate the GCRE as a key regulatory element promoting expression of RITF.

Characterization of Nuclear Proteins from Intestinal Epithelial Cell Lines Binding to GCRE—Using nuclear extracts from various kinds of intestinal cell lines, GMSA was performed to characterize the protein complexes bound to GCRE. To ensure the quality of nuclear protein extracts, GMSA using a consensus Sp1 site as a probe was also performed (Fig. 6A). Crude nuclear protein extracts were prepared from LS174T (5 mg), HT29 (5 mg), undifferentiated (7.5 mg) or differentiated H2 (2.5 mg) subclone of HT29, and Caco2 (2.5 mg) cells, and binding reactions were carried out with the Sp1 or WT2 probe. As sufficient Zn2+ was chelated by the EDTA present in the nuclear extract isolation buffer to prevent Sp1 from binding (30), 1 mM ZnCl2 was added to the GMSA reactions. All of the nuclear proteins bound the Sp1 consensus sequence (Fig. 6A). The identity of the Sp1-binding protein was confirmed by supershift after the addition of anti-Sp1 antibody. These data demonstrated the presence of equal Sp1 binding activity, confirming the adequacy of nuclear extracts from the different cell lines. Using the same reaction conditions, the nuclear proteins from HT29 and Caco2 cells did not appear to bind the GCRE, but strong binding to GCRE was observed by nuclear proteins from LS174T and differentiated H2 cells (Fig. 6B). While little binding of GCRE was observed with nuclear proteins from undifferentiated H2 cells, binding to GCRE was significantly greater by nuclear proteins prepared from H2 cells after goblet cell-like differentiation. These results suggest that expression of nuclear factors that bind to GCRE is associated with differentiation to a goblet cell-like phenotype.

Characterization of Tissue Specificity of the GCRE and Its Binding Protein—GMSA was carried out using nuclear extracts made from cell lines derived from a variety of tissues to further characterize the specificity of expression of the GCRE-binding proteins. As shown in Fig. 7, the GCRE-binding nuclear factor was largely absent in nongoblet cell lines with the exception of HepG2 (liver) and HeLa (cervix) cells. A very small amount of binding protein was also observed in lung (LX-1) cells. Of interest, little binding to GCRE was observed by nuclear proteins from IEC6 cells, a nontransformed rat intestinal crypt cell line. Binding to Sp1 was observed by all of the nuclear proteins in the same fashion as shown in Fig. 6A (data not shown).

To determine whether the GCRE-binding proteins present in extracts from HepG2 and HeLa cells were the same as that in the extract from goblet cells (Fig. 7), transient transfection analysis was undertaken using these cell lines (Fig. 8). The wild-type constructs (1.7WT-luc, 0.2WT-luc, and 0.1WT-luc) and the mutant construct (1.7Δ1-luc) deleted between −154 and −118 were transfected into HepG2 and HeLa cells as well as LS174T cells. Promoter activities were analyzed and expressed as a percentage of the expression of the maximal promoter construct RSV-luc (% RSV-luc) as described under “Materials and Methods” (Fig. 8). Expression of both the truncated...
minimal promoter 0.1WT-luc and the mutant 1.7Δ1-luc were significantly reduced in transfected colonic LS174T cells, compared with expression levels of both the wild-type constructs 1.7WT-luc and 0.2WT-Luc. In contrast, a requirement for the 37-bp sequence was not observed in HepG2 or HeLa cells. Since expression of the truncated minimal promoter 0.1WT-luc in HepG2 and HeLa cells was still high, transient transfection analysis of further truncated constructs (0.08WT-luc and 0.05WT-luc) in these cells was performed (Fig. 8B). Expression of both 0.08WT-luc and 0.05WT-luc was less than 1% of RSV-luc. 0.05WT-luc did not contain TATA box by DNA sequencing, and expression of this construct was not surprisingly the same as the pXP2, the promoterless construct. Transfection analysis using LX-1 cells was also performed, but the expression of all constructs used in this experiment was less than 1% of RSV-luc (data not shown). These results suggest that the protein(s) from HepG2 and HeLa cells that bind the enhancer sequence containing GCRE are distinct from those present in LS174T cells.

Collectively, these studies indicate that a 9-bp element (the

Goblet Cell-specific Enhancer Element in RITF Gene Promoter

FIG. 6. Characterization of Sp1 and GCRE binding by nuclear proteins from intestinal epithelial cell lines. Crude nuclear protein extracts were prepared from various colonic cancer cell lines as indicated: LS174T, HT29, H2 undifferentiated (H2-u), H2 differentiated (H2-d), and Caco2. A, the binding reaction was performed with nuclear protein extracts from each cell line and radiolabeled probe in the absence or presence of unlabeled WT2 competitor. The amounts of nuclear protein were adjusted so that the binding activity in each cell line showed the same strength (LS174T, 5 μg; HT29, 5 μg; H2 undifferentiated, 7.5 μg; H2 differentiated, 2.5 μg; and Caco2, 2.5 μg). a, Sp1; f, free probes. The supershifted complexes by the Sp1 antibody are indicated (s). B, using the same amounts of nuclear protein extracts indicated in A, the reaction was performed with radiolabeled WT2 probe in the absence or presence of unlabeled WT2 competitor. a, binding to WT2; f, free probe.

FIG. 7. Cell specificity of the GCRE-binding protein. GMSA was performed with radiolabeled WT2 probe and crude nuclear extracts made from a variety of cell lines. In addition to LS174T cells, cell lines included the following: TE-1, human esophageal cancer; AGS-B, human gastric cancer; Panc1, human pancreatic cancer; HepG2, human hepatocellular cancer; LLC-PK1, porcine kidney proximal tubule; LX-1, human lung cancer; MCF-7, human mammary adenocarcinoma; HeLa, human cervical adenocarcinoma; and IEC6, rat intestinal crypt epithelial cell. The arrows indicate the position of the delayed complex containing the GCRE-binding protein.

FIG. 8. Expression of RITF promoter constructs in nonintestinal cell lines. HepG2 and HeLa cells were transiently transfected using the calcium phosphate precipitation method as described under Fig. 1 and "Materials and Methods." A, the deletion constructs (1.7WT-, 0.2WT-, and 0.1WT-luc) and a deletion mutant (1.7Δ1-luc) were the same as those shown in Fig. 5. B, further deletion constructs (0.06WT-luc and 0.05WT-luc) were generated by polymerase chain reaction from 1.7WT-luc. Results are expressed as a percentage of RSV promoter-driven luciferase activity. All results are the average of three independent transfections and are expressed as the mean ± S.E. Data from LS174T cells are represented in Fig. 1 and shown for comparison.

GCRE) of −135 and −127 in the RITF promoter is a cis-active element directing goblet cell-specific expression. This element appears to be bound by a distinct nuclear transcription factor present in goblet-like cell lines.
DISCUSSION

Goblet cells are abundant constituents of the surface epithelium within the small and large intestine, but characterization of the molecular basis of goblet cell differentiation and function has been quite limited. Differentiation of epithelial cells in the gastrointestinal tract is a complex and dynamic process. In normal mucosa, not only is the tissue-specific phenotype maintained along the longitudinal axis from esophagus to large bowel, but vertical differentiation from crypt to villus is sustained as well. Moreover, differentiation into several region-specific subpopulations is observed within the epithelium along the length of the gastrointestinal tract. Among the growing list of cloned genes whose products are intestine-specific, TTF represents the first within the gastrointestinal tract exclusively expressed by goblet cells. Following earlier initial studies of molecular cloning of the RITF gene (15), in this paper we report the identification of a goblet cell-specific enhancer element in the RITF gene promoter bound by a goblet cell nuclear protein.

Trefoil proteins, including ITF, are secreted onto the mucosal surface and appear to function to preserve mucosal integrity, protecting the epithelium from injury by a variety of noxious agents (31–33). Furthermore, these factors facilitate rapid healing after injury by promoting restitution, the initial phase of epithelial migration that reestablishes surface continuity (7, 31). Mice rendered deficient in ITF through targeted gene deletion are exquisitely sensitive to injury by standard agents (e.g. dextran sodium sulfate) due to impaired restitution (34). An ulcer-associated cell lineage has been reported to appear adjacent to areas of gastrointestinal ulceration, with cells containing EGF-immunostaining material in the base of the newly budding cell lineage and trefoil protein-producing cells appearing more distally along the developing ductule (35).

There has been limited past characterization of the trefoil gene regulatory elements and as yet no delineation of the regulatory elements that are responsible for the regional selective expression of the different trefoil peptides. The presence of an EGF-responsive element in the 5′-flanking region of the pS2 gene (19) has led to speculation about the role of EGF in inducing expression of trefoil proteins in response to mucosal injury (36). However, scrutiny of the 5′-flanking region of the ITF gene demonstrated no significant homology to known EGF response elements (15).

Transient expression of deletion constructs containing various lengths of RITF gene 5′-flanking region ligated to a luciferase reporter gene indicate that an element present between −154 and −118 of RITF 5′-flanking sequences is able to enhance expression in a goblet cell-specific manner (Fig. 1). Subsequent mutational analysis indicates that this regulatory potential is conferred by a 9-bp element. Moreover, GMSA demonstrates the presence of two proteins that bind specifically to the same sequence (CCCCTCCCC, named GCRE) in nuclear extracts of goblet cells (Figs. 2 and 3). A search of this 5′-flanking region of the RITF gene reveals none of the known regulatory elements that have been demonstrated to play a role in intestine-specific expression. Thus, with the exception of some AT repeats, no areas of significant homology appear to exist within the full-length RITF promoter and the reported 5′-flanking regions of the genes from human intestinal alkaline phosphatase (37), human intestinal fatty acid-binding protein (20), porcine aminopeptidase N (38), human and mouse sucrase-isomaltase (21, 39), or human lactase-phlorizin hydrolase (22). However, there is an Sp1-like motif (CCTCCCC) in the GCRE, and one of the two specific binding proteins appears to be one of the Sp binding proteins, a family of zincfinger proteins (Fig. 4). The Sp1 motif is present in the enhancer regions of diverse genes (40, 41). While it appears that one of the proteins present in nuclear extract from goblet cell-like lines binds in an Sp1-like fashion, it is also apparent that these extracts contain a protein that is distinct from an Sp1 element that binds the GCRE.

The GMSA using nuclear extracts from various undifferentiated H2 cells grown in standard conditions contains little GCRE-binding protein (Fig. 6). In contrast, differentiated H2 cells, which exhibit a goblet cell-like phenotype, contain much greater nuclear protein bound to GCRE than undifferentiated H2 cells. No specific binding reaction to GCRE is observed with the nuclear proteins from Caco2 cells, which exhibit a columnar enterocyte phenotype (Fig. 6). Although the DNA-protein complexes observed in experiments using the nuclear proteins from LS174T and H2 cells are the same in size and although their binding appears to be specific as indicated by competition assays, it remains possible that these GCRE-binding proteins are different. At a minimum, these studies indicate that intestinal cells that have differentiated to goblet cell-like phenotype possess nuclear binding proteins that recognize the GCRE.

The analysis of cell specificity of the GCRE and its binding protein demonstrate that the latter is specifically associated with intestinal goblet-like cells. No GCRE was observed in the nuclear proteins from esophagus, stomach, pancreas, kidney, or breast. While the GCRE-binding proteins are present in liver and cervix cell lines, it appears that this reflects the Sp1-binding protein and another factor that is distinct from that in goblet-like cells because transient transfection assays suggest that the binding of the GCRE sequence in these cells does not promote ITF transcription. Interestingly, IEC6 cells, a non-transformed intestinal epithelial cell line established in this laboratory from neonatal rat intestinal crypt cells (42), do not possess the nuclear factors that bind to the GCRE. Previous studies have shown that IEC6 cells do not have detectable RITF mRNA by Northern blot analysis (14). These findings suggest that some unknown regulatory factor exists for vertical differentiation from crypt to villus in small and large intestine. Transient transfection experiments comparing LS174T cells and cells derived from other organs suggest that the GCRE of the RITF gene promoter is capable of directing intestine and goblet cell-specific expression.

In summary, we have identified a goblet cell-specific enhancer element in RITF gene promoter bound by a goblet cell nuclear protein. Further characterization of this enhancer element may provide insight into the molecular basis of the goblet cell phenotype. Future studies may also identify the genetic elements responsible for “ectopic” expression of ITF in pathologic conditions.

Acknowledgments—We thank Drs. T. C. Wang and H. Nakagawa at Massachusetts General Hospital for helpful discussions.

REFERENCES

1. Thim, L. (1986) Biochem. J. 233, 399
2. Carr, M. D. (1992) Biochemistry 31, 1998–2004
3. Hoffmann, W., and Hauser, F. (1993) Trends Biochem. Sci. 18, 239–243
4. Thim, L. (1989) FEBS Lett. 250, 85–90
5. Tomasetto, C., Rio, M.-C., Gautier, C., Wolf, C., Hareveni, M., Chambon, P., and Lathe, R. (1990) EMBO J. 9, 407–414
6. Hauser, F., and Hoffmann, W. (1992) J. Biol. Chem. 267, 24620–24624
7. Jorgensen, K. H., Thim, L., and Jacobsen, H. E. (1982) Regul. Pept. 3, 207–219
8. Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A., and Chambon, P. (1982) Nucleic Acids Res. 10, 7895–7903
9. Rio, M. C., Belloog, J. P., Daniel, J. Y., Tomasetto, C., Lathe, R., Chened, M. P., Batzschelager, A., and Chambon, P. (1988) Science 241, 705–708
10. Lefebvre, O., Wolf, C., Kedingger, M., Chambon, M. P., Tomasetto, C., Chambon, P., and Rio, M. C. (1993) J. Cell Biol. 122, 191–198
11. Jeffrey, G. P., Oates, P. S., Wang, T. C., Babuyshty, M. W., and Brand, S. J. (1994) Gastroenterology 106, 336–345
12. Podolsky, D. K., Lynch, Devaney, K., Stow, J. L., Oates, P., Murgue, B., Defournment, M., Sands, B. E., and Mahuta, Y. R. (1993) J. Biol. Chem. 268, 6694–6702
13. Seid, T., Dooley, S., and Welter, C. (1995) Biochem. Biophys. Res. Commun. 214, 195–199
14. Suenori, S., Lynch, Devaney, K., and Podolsky, D. K. (1991) Proc. Natl. Acad.
Identification of a Goblet Cell-specific Enhancer Element in the Rat Intestinal Trefoil Factor Gene Promoter Bound by a Goblet Cell Nuclear Protein
Haruhiko Ogata, Nagamu Inoue and Daniel K. Podolsky

J. Biol. Chem. 1998, 273:3060-3067.
doi: 10.1074/jbc.273.5.3060

Access the most updated version of this article at http://www.jbc.org/content/273/5/3060

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 42 references, 16 of which can be accessed free at http://www.jbc.org/content/273/5/3060.full.html#ref-list-1