Transcriptional Regulation of the Human Polymeric Ig Receptor Gene: Analysis of Basal Promoter Elements

Paula M. Hempen, Kimberly M. Phillips, Pamela S. Conway, Katrina H. Sandoval, Tracey A. Schneeman, Hsin-Jung Wu and Charlotte S. Kaetzel

*J Immunol* 2002; 169:1912-1921; doi: 10.4049/jimmunol.169.4.1912
http://www.jimmunol.org/content/169/4/1912

---

**References**

This article *cites 49 articles*, 26 of which you can access for free at:
http://www.jimmunol.org/content/169/4/1912.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Secretory Igs provide the first line of adaptive immune defense against ingested, inhaled, and sexually transmitted pathogens at mucosal surfaces. The polymeric Ig receptor regulates transport of dimeric IgA and pentameric IgM into external secretions. The level of expression of polymeric Ig receptor is controlled to a large extent by transcription of the PIGR gene in mucosal epithelial cells. Here we present a detailed analysis of the promoter of the PIGR gene by transient transfection of luciferase reporter plasmids into cultured cell lines. Comparisons of the human and mouse PIGR promoters in human and mouse intestinal and liver cell lines demonstrated that the human PIGR promoter was 4- to 5-fold more active than the mouse PIGR promoter in all cell types, and that both the human and mouse PIGR promoters were more active in intestinal than in liver cell lines. Targeted deletions of 22-bp segments of the human PIGR promoter revealed that the region from nt −63 to −84 is crucial for basal transcription, and that two upstream regions can act as positive or negative regulators. Point mutations within the region from nt −63 to −84 demonstrated that an E box motif, which binds the basic helix-loop-helix protein upstream stimulatory factor, is required for PIGR promoter activity. Two additional regulatory motifs were identified in the proximal promoter region: a binding site for AP2, and an inverted repeat motif that binds an unidentified protein. These findings suggest that cooperative binding of multiple transcription factors regulates basal activity of the human PIGR promoter.

PIGR promoter is regulated by cooperative interactions among multiple transcription factors.

Materials and Methods

Cell culture

The CaCo2 human colon carcinoma cell line, the HepG2 human hepatoma cell line, and the Hepal-6 mouse hepatoma cell line were obtained from American Type Culture Collection (Manassas, VA). Transmortalized mouse intestinal cells (m-ICcl2) that maintain a crypt phenotype (30) were a gift from Dr. J.-P. Kraehenbuhl (University of Lausanne, Lausanne, Switzerland). CaCo2 cells were maintained in RPMI 1640 medium with a final concentration of 25 mM HEPES, 2 mM L-glutamine, 10% FCS, and antibiotics (100 U/ml penicillin, 100 μg/ml streptomycin, and 250 μg/ml fungizone). HepG2 cells were maintained in DMEM/Ham’s F-12 (1/1) medium with a final concentration of 15 mM HEPES, 2.5 mM t-glutamine, 10% FCS, and antibiotics. Hepal-6 cells were maintained in DMEM with a final concentration of 4.5 g/L glucose, 20 mM HEPES, 4 mM t-glutamine, 10 μg/ml insulin, 10% FCS, and antibiotics. The m-ICcl2 cells were grown in DMEM/Ham’s F-12 (1/1) medium with a final concentration of 20 mM HEPES, 2.5 mM t-glutamine, 10 μg/ml insulin, 5.5 μg/ml transferrin, 6.7 ng/ml sodium selenite, 2% FCS, and antibiotics. All cell culture reagents were obtained from BioWhittaker (Walkersville, MD), except for insulin, transferrin, and sodium selenite (Life Technologies, Gaithersburg, MD). Cells were maintained at 37°C in a humidified 5% CO2 atmosphere.

Construction of chimeric plgR-luciferase reporter plasmids

Chimeric reporter plasmids were constructed in which sequences from the PIGR gene were inserted upstream of the firefly luciferase gene in the pGL2-Basic plasmid (Promega, Madison, WI) as previously described (20). Internal deletions and mutations in the human PIGR promoter were generated by the two-step megaprimer PCR method (31). In the first PCR step, mutagenic primers were paired with a lower primer corresponding to the reverse complement of nt 1 to +29, and a 12-nt extension at the 5′ end that generated a HindIII site. All primers were obtained from Life Technologies. The wild-type (WT) PIGR promoter was used as a template. PCR was performed using the AmpliWax PCR gel-facilitated hot start method (PerkinElmer, Norwalk, CT) with ELOONGASE Enzyme Mix (Life Technologies). The conditions for PCR were 30 cycles of 94°C for 2 min, 55°C for 2.5 min, and 72°C for 4 min, followed by a final extension at 72°C for 10 min. The products of these reactions were gel-purified and used as the lower megaprimers in the second PCR reaction. The upper primer corresponded to nt −280 to −251 and a 12-nt extension at the 5′ end that generated an XhoI site. The amplified products from the second PCR reactions were gel-purified, digested with XhoI and HindIII, and cloned into the promoterless pGL2-Basic plasmid. The sequences of all plasmids were verified by automated fluorescence sequencing.

Transfection of cells and analysis of PIGR promoter activity

Cells were transiently cotransfected with reporter-luciferase plasmids and a control vector in which the CMV immediate early promoter controls expression of the bacterial chloramphenicol acetyltransferase (CAT) gene (pCDNA3.1(−)CAT; Invitrogen, San Diego, CA; Figs. 2 and 3) or Renilla luciferase gene (pRL-CMV; Promega; Figs. 4 and 6). Cells were plated in 35-mm culture wells at ~70–80% confluence 24 h before transfection. Two micrograms of PIGR promoter-luciferase plasmid and either 1 μg pCDNA3.1(−)CAT plasmid or 0.4 μg pRL-CMV plasmid were mixed with Tfx-50 reagent (Promega; 2:1 Tfx-50:DNA charge ratio) in a final volume of 1 ml serum-free medium and incubated with cells for 4 h at 37°C. Cells were then overlaid with 2 ml complete medium with serum, without removing the Tfx-50:DNA mixture. Culture media were changed daily until the cells were harvested at 72 h (Figs. 2–4) or 24 h (Fig. 5) after transfection. For the experiments shown in Figs. 2 and 3, preparation of cell lysates and analysis of luciferase and CAT activities were performed as previously described (20). For the experiments shown in Figs. 4 and 5, firefly and Renilla luciferase activities were analyzed using the Dual-Luciferase Reporter Assay System according to the manufacturer’s protocol (Promega). Promoter activity was calculated as [firefly luciferase activity/ control enzyme activity (CAT or Renilla luciferase)] × 100%.

Preparation of nuclear extracts and EMSA

Nuclear protein extracts were prepared as previously described (32), and protein concentrations were determined by the Bradford assay (33) (Bio-Rad, Hercules, CA). EMSAs were performed as previously described (20) with the following modification to reduce the formation of nonspecific protein/DNA complexes. Nuclear extracts (10 μg protein) were preincubated for 15 min at room temperature with 2 μg poly dI-C') plus unlabeled homologous or heterologous double-stranded oligodeoxynucleotides. Next, 10–14 fmol 32P-labeled double-stranded oligodeoxynucleotide was added, and incubation was performed for an additional 15 min at room temperature. Where indicated, 2 μg Abs to the transcription factors USF1, USF2, c-Myc, Max, or IRF1 (Santa Cruz Biotechnology, Santa Cruz, CA) were added and incubated for 1 h at 5°C. Protein-bound DNA was separated from free DNA by nondenaturing PAGE as previously described (20).

DNase I footprinting analysis

A nonradioactive detection method was developed by end-labeling dsDNA probes with digoxigenin (DIG), and detection with anti-DIG Abs. DNA probes, spanning nt −280/−47 or −280/−29 of the PIGR gene, were generated by PCR as described above, except that the upper primer was labeled at its 5′ end with DIG according to the manufacturer’s protocol (Roche, Indianapolis, IN). Purified DIG-labeled DNA was incubated for 10 min on ice with or without added protein (50 μg nuclear protein from CaCo2 or HepG2 cells, or increasing volumes of bacterial extract containing 0.5, 1.0, or 2.0 μg recombinant AP2 protein). DNase I digestions were conducted with the Core Footprinting System (Promega). The concentrations of DNase I and the incubation times were optimized for each sample and are listed in the figure legends. We determined in preliminary experiments that it was necessary to add higher concentrations of DNase I in the presence of nuclear extracts from CaCo2 and HepG2 cells to achieve a degree of digestion of the DNA probe equivalent to that seen in the absence of extract. DNA sequencing ladders were generated using the T7 DNA polymerase chain termination method (Sequenase version 2.0; Amersham Life Science, Cleveland, OH) and the same 5′−DIG-labeled oligodeoxynucleotide that was used to generate the labeled footprinting probes. The DNase I-digested samples and the sequencing ladders were separated by electrophoresis in a 6% polyacrylamide gel containing 100 mM Tris-borate (pH 8.3), 1 mM EDTA, and 7.0 M urea. The gel was capillary blotted overnight onto a charged nylon membrane (Osmonics Laboratory Products, Minnetonka, MN) and baked for 1 h at 80°C. DIG-labeled DNA was detected using anti-DIG Ab and an alkaline phosphatase-based fluorescent detection system (Roche). Fluorescent membranes were exposed at room temperature to XAR AF 1417 film (Action Scientific, Carolina Beach, NC).

Data analysis

Analyses of free energy changes associated with DNA stem-loop formation were performed using OLIGO 4.0 Primer Analysis Software (National Bio-sciences, Plymouth, MN). Two-way ANOVA was used to test for statistical differences in promoter activity, controlling for multiple independent experiments (StatView; SAS Institute, Cary, NC). For Figs. 2 and 5, individual comparisons were made using Student’s t test. For Figs. 3 and 4, pairwise multiple comparisons were made using Fisher’s protected least significance difference test.

Results

Tissue- and species-related effects on the activity of the PIGR promoter

We previously reported that the region from −280/+29 of the human PIGR gene was sufficient to promote transcription of a heterologous reporter gene in a human colon carcinoma cell line (20). Fig. 1 illustrates an alignment of this region of the human PIGR gene with the corresponding regions of the rat and mouse PIGR genes. The region from −206/+29 of the human PIGR gene has overall similarities of 63 and 66% to the region from −188/−25 of the rat and mouse PIGR genes, respectively. Upstream of position −206, however, the sequence of the human PIGR gene diverges completely from the rodent PIGR genes. To compare directly the activities of the human and mouse PIGR promoters, the homologous regions were cloned into reporter vectors containing the firefly luciferase gene, which were transiently transfected into human and mouse intestinal and liver epithelial cell lines (Fig. 2). We chose intestine and liver because the endogenous expression of plgR mRNA varies between these tissues in a species-specific pattern (see Discussion). The human PIGR promoter was significantly more active than the mouse homolog in all cell lines tested regardless of the tissue of origin: 4-fold higher in CaCo2 human colon...
carcinoma cells, 6-fold higher in HepG2 human hepatoma cells, 4-fold higher in m-IC\textsubscript{12} mouse intestinal epithelial cells, and 9-fold higher in Hepa1\textsubscript{6} mouse hepatoma cells. In addition, \textit{PIGR} promoter activity was consistently higher in intestinal than in liver cell lines. The human \textit{PIGR} promoter was 11-fold more active in human intestinal cells than in human liver cells and 7-fold more active in mouse intestinal cells than in mouse liver cells. Similarly, the activity of the mouse \textit{PIGR} promoter was 16-fold higher in human intestinal cells than in human liver cells and 15-fold higher in mouse intestinal cells than in mouse liver cells. Although we observed a consistent pattern, we caution that our conclusions are based on results from only four cell lines. It should also be noted that \textit{PIGR} promoter activity was expressed relative to the activity of a cotransfected CMV promoter. We cannot rule out the possibility that tissue differences in CMV promoter activity contributed to the observed differences in \textit{PIGR} promoter activity between intestinal and liver cell lines.

Analysis of potential regulatory regions in the human \textit{PIGR} promoter

Comparison of the sequences of the human \textit{PIGR} promoter within the region from \{-266\}+1 with the corresponding regions of the rodent \textit{PIGR} promoters revealed subregions of significant similarity or divergence (Fig. 1). We designated four subregions (sites A–D) in which notable differences between the human and rodent \textit{PIGR} genes were observed. Sites A and B differ significantly in sequence between the human and rodent \textit{PIGR} genes, and site C contains an extra 18-bp sequence with no homolog in the rodent \textit{PIGR} genes. Homology between the human and rodent \textit{PIGR} genes is higher in site D, except for a critical A-G difference within the E box motif and differences in the nucleotides flanking the 3' end of the E box. At positions \{-120\} (between sites B and C) and \{-63\} (downstream of site D) are two 10-bp motifs that are complementary in sequence (i.e., inverted repeats). The rodent \textit{PIGR} genes also contain a 10-bp inverted repeat at this position, which are 90% identical in sequence to the inverted repeats in the human \textit{PIGR} gene. A consensus binding site for the transcription factor AP2 was observed at position \{-48\} of the human, but not the rodent, \textit{PIGR} promoters. To determine which subregions of the human \textit{PIGR} promoter bind nuclear proteins, we performed a DNase I footprint analysis of the region spanning nt \{-280\} (Fig. 3). Binding of nuclear proteins from CaCo2 cells completely protected the region spanning site D and the second inverted repeat from DNase I digestion. Binding of nuclear proteins from HepG2 cells caused a similar, but less striking, footprint over site D, consistent with the lower promoter activity in HepG2 vs CaCo2 cells (Fig. 2). Binding of nuclear proteins from CaCo2 cells, but not HepG2 cells, created a DNase I-hypersensitive site at the junction of sites C and D. This hypersensitive site may have resulted from bending of the DNA strand induced by binding of proteins to site D. While binding of nuclear proteins caused some minor changes in DNase I sensitivity in sites A, B and C, they were not as striking as the changes in site D. These results suggest that recruitment of nuclear proteins to site D may be important for transcription of the \textit{PIGR} gene.
FIGURE 2. Tissue- and species-related effects on the activity of the PIGR promoter. Reporter plasmids were constructed in which the human PIGR promoter (−210/+29) or the corresponding region of the mouse PIGR promoter (−188/+25) were cloned upstream of the luciferase gene in the pGL2-Basic plasmid. The designation Null refers to the promoterless pGL2-Basic plasmid. The reporter-luciferase plasmids along with the pcDNA3.1(−)CAT plasmid to control for transfection efficiency were transiently transfected into the following human or mouse epithelial cell lines: top left, CaCo2 human colon carcinoma; top right, HepG2 human hepatoma; bottom left, m-IC, mouse enterocyte; and bottom right, Hepa1–6 mouse hepatoma. Cell lysates were analyzed for promoter activity (luciferase/CAT) as described in Materials and Methods. Data are reported as the mean ± SEM (n = 6). a, Mean is significantly less than the mean for the human PIGR promoter plasmid (p < 0.001); b, mean is significantly greater than the mean for the null plasmid (p < 0.05).

Since recruitment of nuclear proteins to the PIGR promoter was striking in CaCo2 cells, we used that cell line to test the effects of specific mutations on promoter activity. We began by creating internal deletions of sites A, B, C, and D (Fig. 3B). All deletions were 22 bp in length, representing approximately two turns of the DNA helix, to minimize potential effects of disrupting spacing between regulatory elements. This approach has been used to preserve the stereospecific alignment of regulatory elements in the promoter of the HLA-DRA gene (34). Deletion of site A caused a modest, but highly reproducible, decrease in promoter activity. Deletion of site B had no effect, while deletion of site C caused a significant increase in promoter activity. Deletion of site D caused a dramatic decrease in PIGR promoter activity to a level that was not significantly different from that of the promoterless luciferase plasmid. These results suggested that sites A and D may contain positive regulatory elements, and that an element(s) in site D may be required for basal transcription of the PIGR gene. In contrast, an element(s) in site C may act as a negative regulator(s) of PIGR promoter activity.

Effects of point mutations in potential regulatory motifs on activity of the human PIGR promoter

The identification of site D as a major focus of protein binding and promoter activity led us to investigate the E box motif at position −71. The canonical E box sequence CACGTG was changed to the canonical E box sequence CACCTG (nucleotides differing from the WT human PIGR sequence are underlined). For the E box up mutation, the core sequence CACCTG was changed to CACAGA, which has been reported to abolish binding of bHLH-zip proteins (36). We also included the mouse PIGR promoter in this analysis, since its E box has the canonical CACCTG sequence (see Fig. 1). The E box up mutation caused a significant increase in PIGR promoter activity, presumably due to enhanced binding of bHLH-zip proteins. In contrast, the E box down mutation caused a significant decrease in PIGR promoter activity to a level that was statistically indistinguishable from the promoterless control plasmid. This effect was comparable to deletion of the entire site D (Fig. 3B), confirming our hypothesis that the E box is responsible for the transcriptional activity of site D. As shown in Fig. 2, the activity of the mouse PIGR promoter was significantly less than that of the human PIGR promoter despite the fact that the mouse PIGR promoter contains the canonical E box sequence (based on the multiple group comparison test used to analyze the data in Fig. 4A, the activity of the mouse PIGR promoter appeared to be indistinguishable from the promoterless control plasmid; however, an individual comparison by t test did detect a significant difference between the mouse PIGR promoter and the control plasmid). This result suggests that sequences outside the E box are responsible for the enhanced activity of the human PIGR promoter relative to its murine counterpart.

We next tested the effects of mutations in the 10-bp inverted repeats at positions −120 and −63 of the human PIGR promoter. The sequence of repeat 2 is the reverse complement of the sequence of repeat 1, raising the possibility that stem-loop structures could form by base-pairing of the upper and lower strands of repeat 1 with the corresponding strands of repeat 2 (Fig. 4B). Base-pairing of these 10-nt sequences would theoretically be associated with a ΔG of −11.9 kcal/mol. Formation of stem-loops could result in transcriptional repression, since the important E box motif between repeat 1 and repeat 2 would be disrupted. Conversely, binding of bHLH-zip proteins such as USF to the E box motif should stabilize the DNA in a double-stranded conformation, inhibiting stem-loop formation. To test the hypothesis that stem-loop formation may regulate the activity of the PIGR promoter, we created a series of mutations in the two inverted repeats, as described in Fig. 1. For MutRpt1 and MutRpt2, multiple point mutations were made in repeats 1 and 2, respectively, while the opposite repeat was maintained in the WT configuration. Either of these mutations would prevent stem-loop formation. The mutations in repeats 1 and 2 were designed to be complementary, so that stem-loop formation could be restored in the double mutant (predicted ΔG, −9.2 kcal/mol). However, if the actual sequences of the repeats were important (for example, for DNA-protein interactions), then the double mutant promoter might have altered activity relative to the WT promoter. Mutations in repeat 1 caused a 50% increase in PIGR promoter activity, consistent with the hypothesis that disruption of stem-loop formation could enhance promoter function (Fig. 4A). The relatively small difference in activity between the MutRpt1 and WT promoters could have been because binding of protein(s) to the E box constitutively inhibited stem-loop formation in the WT promoter. If this were the case, then the reduced activity of the E box down mutant could have partly been due to enhanced stem-loop formation secondary to loss of protein binding at the E box. To test this hypothesis we created a new reporter plasmid that contained both the MutRpt1 and E box down mutations (MR1/EB Dn), in which disruption of stem-loop formation might compensate for loss of E box function. However,
we found that the promoter activity of the combined MR1/EB Dn mutant was just as low as that of the single E box down mutant (Fig. 4A). We subsequently discovered that while the E box down mutation specifically inhibited binding of the transcription factor USF to the E box, other unidentified proteins remained bound (see Fig. 6). We therefore favor the hypothesis that the reduced promoter activity of the E box down mutant is due to changes in the composition of proteins bound to the E box and not to enhanced stem-loop formation.

In contrast to repeat 1, mutations in repeat 2 caused a significant decrease in PIGR promoter activity to a level indistinguishable from that of the promoterless control vector. Activity of the double mutant promoter was the same as that of the MutRpt2 promoter, suggesting that decreases in PIGR promoter activity were due to mutations in the sequence of repeat 2 and not to changes in stem-loop formation. In this context it is significant that nuclear extracts from CaCo2 cells protected the region around repeat 2, but not repeat 1, from DNase I digestion (Fig. 3A). Repeat 2 may therefore

**FIGURE 3.** Protein binding and transcriptional activity of sites A–D of the human PIGR promoter. A, DNase I footprinting analysis. A dsDNA probe spanning nucleotides −230 to −47 of the human PIGR gene was labeled at the 5′ end of the sense strand and digested for 90 s with DNase I in the presence or the absence of 50 μg nuclear protein from CaCo2 cells or HepG2 cells. The amounts of DNase I added were (left to right) 0.05, 0.1, 0.2, 0.3, 1.2, 1.4, 1.2, and 1.4 U. The lanes labeled G, A, T, and C represent a DNA sequencing ladder analyzed alongside the DNase I-digested probe. ●, regions of the DNA probe that were protected by binding of nuclear proteins. The arrow represents a DNase I-hypersensitive site at the 5′ end of site D. B, Effects of deletions of sites A–D on PIGR promoter activity. Mutant promoters containing 22-bp internal deletions are designated ΔA, ΔB, ΔC, and ΔD. The reporter-luciferase plasmids and the pcDNA3.1(−)/CAT control plasmid were transiently transfected into CaCo2 cells, and cell lysates were analyzed for luciferase/CAT activity. Data from three independent experiments were combined (n = 18) and normalized to the activity of the WT PIGR promoter. Data are reported as the mean ± SEM. a, Mean is significantly different from the mean for the WT PIGR promoter plasmid (p < 0.05); b, mean is significantly different from the mean for the null plasmid (p < 0.05).

**FIGURE 4.** Effects of mutations in the E box and inverted repeats on PIGR promoter activity. A, Reporter-luciferase plasmids containing the indicated mutations (see Fig. 1) were transiently transfected with the pRL-CMV control plasmid into CaCo2 cells, and cell lysates were analyzed for firefly/Renilla luciferase activities. Data from three independent experiments were combined (n = 18) and normalized to the activity of the WT human PIGR promoter. Data are reported as the mean ± SEM. a, Mean is significantly different from the mean for the WT human PIGR promoter plasmid (p < 0.05); b, mean is significantly different from the mean for the null plasmid (p < 0.05). B, Theoretical stem-loop formation created by base-pairing of repeat 1 with repeat 2.
act as a positive regulatory motif by associating with a specific transcription factor(s). Because of the strategic location of Repeat 2 between the E box and a potential binding site for AP2 (see Fig. 1), it was important to determine whether the mutations we created in repeat 2 had disrupted the activity of adjacent regulatory elements. For example, the mutations in repeat 2 included an A-T transition at the nt immediately flanking the 3′ end of the E box, a position that can affect binding of transcription factors to the E box (36). To isolate the effect of this mutation, we created a single point mutation at this position (E box flank A-T; see Fig. 1). This mutation had no effect on PIGR promoter activity (Fig. 4A), indicating that the suppressive effect of the mutations in repeat 2 was not due to disruption of E box function. We next used DNase I footprinting analysis to demonstrate that the mutations in repeat 2 did not inhibit binding of recombinant AP2 to the adjacent motif (Fig. 5A). To determine whether the AP2 site was active in CaCo2 cells, we tested the effects of mutations in the AP2 site on PIGR promoter function. The CC dinucleotide at position −46 was the location of the major AP2 footprint (see arrows in Figs. 1 and 5A). We changed the C at nt 66 to an A to match the sequence of the rodent PIGR genes, which do not have a consensus AP2 motif at this location. We also mutated the GG dinucleotide flanking the 5′ end of the AP2 site to CT, as found in the rodent PIGR genes. These mutations disrupted binding of recombinant AP2 to the PIGR promoter (Fig. 5A) and significantly reduced PIGR promoter activity (Fig. 5B). Taken together, these results demonstrate that mutations in the repeat 2 motif did not affect the function of the adjacent E box and AP2 elements and, alternatively, may have inhibited binding of a unique protein(s) to this region of the PIGR promoter. We have not been able to identify protein(s) binding to the either the repeat 1 or 2 motifs by EMSA (data not shown), but this does not rule out the possibility that binding of transcription factor(s) to repeat 2 may require complex interactions with factors bound to the adjacent E box and AP2 motifs.

**Identification of the major E box binding protein in nuclear extracts of CaCo2 cells**

To determine the sequence requirements for binding of nuclear proteins to site D of the human PIGR promoter, we performed EMSAs using as probe a WT site D oligo or an oligo containing the E box up mutation. Three specific DNA-protein complexes were observed in nuclear extracts from CaCo2 cells (Fig. 6B). The pattern of DNA-protein complexes was similar for the site D and E box up probes, although the intensities of the bands were stronger with the E box up probe. The pattern of competition by unlabeled DNA revealed a higher binding affinity for the E box up oligo. Weak displacement was seen with a 200-fold molar excess of unlabeled site D oligo, while the same concentration of E box up oligo completely displaced complexes 1 and 2 (but not complex complex...
3). Surprisingly, the E box down oligo caused some displacement of complex 1 and complete displacement of complexes 2 and 3 from the WT site D oligo. However, this oligo did not compete with the higher affinity E box up oligo. The simplest explanation for the partial competition by the E box down oligo is that this mutation lowered the affinity, but did not eliminate binding of transcription factors to the E box motif. Some competition was therefore observed with the low affinity WT E box probe, but not with the high affinity E box up probe, in an in vitro binding assay. In living cells, however, the lower affinity of the E box down motif was insufficient to support PIGR promoter activity (Fig. 4A). To test whether binding of proteins to site D required the E box motif, we used a competitor oligo (Myc) that was homologous to site D only within the E box and which contained the higher affinity CACGTG motif. At a 200-fold molar excess, the Myc oligo displaced all three specific protein complexes from the WT site D probe, confirming that binding of all three complexes required the E box motif. These results were confirmed by testing a range of concentrations of site D and Myc oligo competitors (data not shown).

To identify the proteins binding to site D, we performed an EMSA in which specific DNA-protein complexes were supershifted by Abs to known members of the bHLH-zip family of transcription factors. We first tested Abs to USF1 and USF2, protein subunits that combine as hetero- and homodimers to form the transcription factor, USF (37). Using the WT site D probe, Abs to USF1 and USF2 strongly inhibited the formation of DNA-protein complex 1 and also caused a weak supershift (Fig. 6A). Using the higher affinity E box up probe, a stronger supershift was observed. Abs to USF1 completely supershifted complex 1, while Abs to USF2 supershifted most, but not all, of complex 1. No supershift was observed with Abs to IRF1, an unrelated transcription factor. These results are consistent with the identification of USF as the protein constituent of complex 1. The USF appears to be comprised predominantly of heterodimers of USF1 and USF2, which are supershifted by Abs to either USF1 and USF2, and a lesser amount of USF1 homodimers, which are supershifted by Abs to USF1, but not USF2. These results suggest that the increased activity of the E box up and decreased activity of the E box down promoters was due to increases and decreases in binding of USF. Abs to e-Myc and its binding partner Max did not supershift any of the site D complexes (data not shown). However, since Myc/Max dimers are inherently unstable and difficult to isolate from cells (38), these results of in vitro binding assays do not necessarily preclude the binding of Myc/Max to site D of the human PIGR promoter in vivo.

**Discussion**

The magnitude of plgR-mediated transport of IgA in the human intestine is impressive; it has been estimated that 3 g IgA is transported daily in the average adult (39, 40). Because a molecule of plgR is consumed with each round of IgA transport, a high rate of synthesis of plgR must be maintained by intestinal epithelial cells. To begin to understand mechanisms of PIGR gene regulation, we have analyzed PIGR promoter activity in human and mouse intestinal and liver cell lines. Our data suggest a model through which a central E box enhancer element cooperates with multiple other elements to regulate transcription of the PIGR gene.

A direct comparison of the homologous regions of the human and mouse PIGR promoters demonstrated that the transcriptional...
activity of the human \textit{PIGR} promoter was significantly higher in both human and mouse cell lines (Fig. 2). Subsequent analyses did not define a specific region unique to the human \textit{PIGR} gene that would explain the species differences in \textit{PIGR} promoter activity. We therefore favor a model by which complex interactions among multiple regulatory sites contribute to the superior activity of the human \textit{PIGR} promoter. The tissue specificity of plgR mRNA expression varies between humans and rodents (3). Northern blot analysis of human tissues demonstrated that plgR mRNA was expressed at high levels throughout the small and large bowels, but was undetectable in the liver (41). By contrast, Northern blot analysis of mouse tissues demonstrated high expression of plgR mRNA in both intestine and liver (42). A quantitative RT-PCR analysis of mouse tissues indicated that expression of plgR mRNA (relative to \( \beta \)-actin mRNA) was 2- to 3-fold higher in intestine than in liver (25). In the present study the activities of both the human and mouse \textit{PIGR} promoters were 5- to 10-fold higher in intestinal than in liver cell lines, suggesting that tissue-specific mechanisms may contribute to the regulation of \textit{PIGR} transcription. The higher activity of the mouse \textit{PIGR} promoter in intestinal than in liver cell lines is consistent with the higher endogenous expression of plgR mRNA in mouse intestine. Although the activity of the human \textit{PIGR} promoter was higher in intestinal than in liver cell lines, our finding of significant human \textit{PIGR} promoter activity in liver cell lines is inconsistent with the lack of expression of plgR mRNA in human hepatocytes. It is possible that additional mechanisms may down-regulate steady state plgR mRNA levels in human liver, such as the effects of more distal regulatory regions on \textit{PIGR} gene transcription, liver-specific control of plgR mRNA stability, etc.

The regions designated sites A, B, and C in Fig. 1 represent the most divergent regions between the human and rodent \textit{PIGR} promoters. Deletion of site A caused a modest, but reproducible, decrease in \textit{PIGR} promoter activity in both intestinal and nonintestinal cell lines (Fig. 3), suggesting that it may act as a positive regulatory element. By EMSA, we have detected ubiquitous and intestine-specific proteins that bind to the unique sequence of the human \textit{PIGR} site A region (unpublished observations). We have further observed that point mutations in site A that reduce protein binding also reduce \textit{PIGR} promoter activity (unpublished observations). We have not yet confirmed the identities of the site A binding proteins, and the sequence of site A does not contain binding sites for any known transcription factors in the TRANSFAC database (43). In contrast to site A, site C appears to act as a negative regulatory element (Fig. 3). The sequences of sites B and C are highly related, suggesting that site C may have evolved by duplication and divergence of site B. Both sites contain the motif A\textsc{aaa}gaa\textsc{aaa}nnnaa, which represents a potential binding site for proteins of the IRF family of transcription factors (20). Despite the similarity to IFN-stimulated response elements, sites B and C appear to bind the same set of proteins in the presence and absence of IFN-\( \gamma \) stimulation (20; in Fig. 5 of that reference, site B is designated upstream ISRE-1, site C is designated upstream ISRE-2, and site A is the irrelevant oligo). We are currently characterizing interactions among proteins that bind to sites A, B, and C.

Since deletion of site D completely abolished the activity of the human \textit{PIGR} promoter, we hypothesize that proteins bound to this region are critical for initiation of transcription. Within this region is an E box motif, a consensus binding site for transcription factors of the bHLH-zip family (35). It was previously reported that point mutations in the E box motif caused a significant decrease in activity of the human (22) and mouse (24) \textit{PIGR} promoters. Furthermore, overexpression of USF1, a ubiquitously expressed protein of the bHLH-zip family, resulted in a modest enhancement of the activity of the mouse \textit{PIGR} promoter (24). Here we demonstrated that an abundant nuclear protein complex associated with the \textit{PIGR} E box in CaCo2 cells was a heterodimer of USF1 and USF2, with a minor contribution by USF1 homodimers. Given the critical role of USF in regulation of \textit{PIGR} promoter activity, it is informative to examine the sequence of the \textit{PIGR} E box in the context of known requirements for USF binding. Three approaches have been used to determine the optimal DNA binding site for USF: analysis of USF binding to E box motifs of varying sequence within active promoters (35), analysis of USF binding to oligonucleotides of varying sequence by EMSA (36), and structural analysis of purified USF bound to its cognate DNA element (44). USF shares with all members of the bHLH-zip family an affinity for the core palindromic sequence CACGTG. The DNA-binding surface of USF extends to contact 2 additional bp on either side of the core E box, with a preference for maintaining the palindromic structure across the entire 10 bp. While the prototype sequence for bHLH-zip proteins is RYCA\textsc{gtg}R\textsc{g}, USF shows a clear preference for a T residue in the second position, thus having an optimal binding site of RTCA\textsc{gg}TGA\textsc{g} (36). The sequence of the human \textit{PIGR} E box is a near-perfect match for the USF consensus, with the single exception of the A in position 4 of the core hexanucleotide. Significantly, the DNA binding domain of USF has been shown to contact the core sequence CACATG with a small reduction in binding affinity compared with the palindromic sequence CAC\textsc{gg}T (36). We demonstrated that mutation of the human \textit{PIGR} E box to the canonical sequence CAC\textsc{gg}T significantly increased promoter activity (Fig. 4). We hypothesize that the increase in promoter activity was due to enhanced binding of USF, as supported by our EMSA data (Fig. 6).

In addition to regulating the affinity of USF binding, the sequence of the E box motif in the human \textit{PIGR} promoter may influence the availability of this site for binding other members of the bHLH-zip family of transcription factors. Substitution of an A in position 4 of the core E box motif (CACATG) has been shown to significantly diminish the affinity of Myc/Max binding in vitro while causing only a minimal decrease in the affinity of USF binding (36). Similarly, the presence of a T immediately 5' to the core E box (TC\textsc{cat}AG) favored USF binding while diminishing Myc/Max binding. Similar sequence variations have been shown to affect USF vs Myc/Max binding to the E box motif in the cad gene promoter in living cells (45, 46). Significantly, the human \textit{PIGR} E box contains the sequence TC\textsc{catAG}, which would be predicted to favor USF over Myc/Max binding. We have recently shown that overexpression of USF1 or USF2 enhances and c-Myc inhibits \textit{PIGR} promoter activity in intestinal epithelial cells (M. E. C. Bruno and C. S. Kaetzel, manuscript in preparation). This situation is reversed in the case of the \textit{cad} gene promoter, which is enhanced by c-Myc and inhibited by USF (45, 46). The significance of the opposing effects of USF and c-Myc may relate to their biological roles. USF1 and USF2 have been shown to enhance transcription of genes associated with differentiated cellular function (37), while c-Myc has been shown to enhance the transcription of genes associated with cellular proliferation and oncogenesis, such as the \textit{cad} gene (35). In this context it is significant that plgR expression is correlated with differentiation of intestinal epithelial cells and is diminished in disease states associated with proliferation of intestinal epithelial cells, such as colon cancer and inflammatory bowel disease (3).

We identified a consensus binding site for the AP2 transcription factor that is unique to the human \textit{PIGR} promoter (Figs. 1 and 5). This element bound recombinant AP2, and mutations that disrupted AP2 binding caused a significant decrease in \textit{PIGR} promoter activity. AP2
expression is known to be induced by several signaling pathways, including protein kinase C, cAMP, and retinoic acid (47–49). AP2 may therefore play a role in both basal and inducible transcription of the human PIGR gene.

Positive regulation of the PIGR gene by USF may be relevant to the function of the inverted repeat elements that are highly conserved among the human and rodent PIGR genes (Fig. 1). We hypothesized that the inverted repeats could base pair to form a double stem-loop conformation of the DNA in the proximal promoter region (Fig. 8A). A similar double stem-loop structure has been reported to inhibit transcription of the human platelet thromboxane receptor gene (50). In that system binding of the Sp1 transcription factor to the region between the inverted repeats enhanced transcription by stabilizing the DNA in a double-stranded configuration and inhibiting stem-loop formation. In the case of the PIGR gene we hypothesize that binding of USF to the E box motif, which is located between the inverted repeats, would inhibit stem-loop formation. Given the high level of USF binding activity in CaCo2 cells (Fig. 6), we predict that stem-loop formation would be inhibited strongly. This model explains the modest effect of mutations in repeat 1 on PIGR transcription (Fig. 8), since stem-loop formation would already be inhibited by USF binding to the region between the inverted repeats. Thus, USF may play two roles in regulating the PIGR promoter: first, by acting as a general trans-activating factor; and second, by disrupting the formation of a potentially inhibitory stem-loop structure. The possibility remains that the inverted repeats may play a role in inhibiting PIGR promoter activity in cells where USF activity is low. An alternative role for repeat 2 could involve binding of transcription factor(s).

Mutation of the second inverted repeat caused a dramatic decrease in PIGR promoter activity, which was not due to disruptions in the adjacent E box and AP2 motifs (Figs. 4 and 5). DNase I footprinting analysis indicated that nuclear proteins from CaCo2 cells bound to the repeat 2, but not the repeat 1 motif (Fig. 3). Because the sequences of repeats 1 and 2 are the same (in opposite orientations), we conclude that binding of proteins to these motifs is dependent on context, i.e., interactions with adjacent elements. The repeat 2 motif is strategically placed between the E box and AP2. Significantly, Johansen et al. (22) identified several protein complexes from epithelial cell-lines that bound to a shorter DNA probe including repeat 2, but not the E box, and may represent a unique factor that binds to repeat 2 in cooperation with USF or other E box binding proteins.

In summary, our data suggest a model for transcriptional regulation of the human PIGR gene that involves cooperative binding among multiple transcription factors. Factors that regulate basal transcription may include USF, AP2, a protein(s) that binds the repeat 2 element, and proteins that associate with the regions we have designated sites A and C. During immune and inflammatory responses, these basal regulatory factors may interact with inducible factors such as IRF-1, NF-κB, and STAT6 to up-regulate PIGR gene transcription.

Acknowledgments
We thank Colleen Brady and Andrea McCool for excellent technical assistance.

References


