Cloning and Characterization of the 5'-Transcriptional Regulatory Region of the Human Intercellular Adhesion Molecule 1 Gene*

Klaus Degitz, Li Lian-Jie, and S. Wright Caughman†

From the Department of Dermatology, Emory University School of Medicine, Atlanta, Georgia 30322 and the Dermatology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Cell-cell adhesion is critical in the generation of immunologic responses and is dependent upon expression of a variety of cell surface receptors. While intercellular adhesion molecule 1 (ICAM-1), a specific receptor for lymphocyte function-associated antigen 1, is constitutively expressed by some cell types, its de novo or increased expression by various cells has been associated with the initiation of inflammatory responses and appears to be transcriptionally regulated. The 5' region of the human ICAM-1 gene has been cloned and both structurally and functionally analyzed. A 17.3-kilobase genomic clone containing three exonal regions encoding the N-terminal third of the ICAM-1 protein was isolated. A 2.05-kilobase subclone, containing the 5' most exon, was utilized to determine an interferon-γ-induced transcription initiation site via primer extension and S1 nuclease protection assays. Analysis of the 5'-flanking region revealed consensus sequences for appropriately located basal promoter elements, as well as numerous potential cis-acting enhancer elements. When subcloned into a reporter gene construct, the putative promoter subregion functioned as a potent promoter. However, in accord with biologically observed expression of ICAM-1 in specific cell types, when additional 5'-flanking sequences were included in reporter gene constructs, tissue appropriate repression of transcription was observed.

Intercellular adhesion molecule 1 (ICAM-1) is one of several recently described human cell adhesion molecules that play critical roles in the generation of immunologic responses and is dependent upon expression of a variety of cell surface receptors. While intercellular adhesion molecule 1 (ICAM-1), a specific receptor for lymphocyte function-associated antigen 1, is constitutively expressed by some cell types, its de novo or increased expression by various cells has been associated with the initiation of inflammatory responses and appears to be transcriptionally regulated. The 5' region of the human ICAM-1 gene has been cloned and both structurally and functionally analyzed. A 17.3-kilobase genomic clone containing three exonal regions encoding the N-terminal third of the ICAM-1 protein was isolated. A 2.05-kilobase subclone, containing the 5' most exon, was utilized to determine an interferon-γ-induced transcription initiation site via primer extension and S1 nuclease protection assays. Analysis of the 5'-flanking region revealed consensus sequences for appropriately located basal promoter elements, as well as numerous potential cis-acting enhancer elements. When subcloned into a reporter gene construct, the putative promoter subregion functioned as a potent promoter. However, in accord with biologically observed expression of ICAM-1 in specific cell types, when additional 5'-flanking sequences were included in reporter gene constructs, tissue appropriate repression of transcription was observed.

* This work was supported in part by a grant from the Deutsche Forschungsgemeinschaft (to K. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This work was supported in part by a grant from the Deutsche Forschungsgemeinschaft (to K. D.). The costs of publication of this article were defrayed in part by the payment of page charges. The costs of publication of this article were defrayed in part by the payment of page charges. The costs of publication of this article were defrayed in part by the payment of page charges.

† To whom correspondence should be addressed: Dept. of Dermatology, Emory University School of Medicine, 201 Woodruff Memorial Blvd., Atlanta GA 30322.

The abbreviations used are: ICAM-1, intercellular adhesion molecule 1; kb, kilobase(s); IFN-γ, interferon-γ; HK, human keratinocytes; UTR, untranslated region; CAT, chloramphenicol acetyltransferase; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PIPES, 1,4-piperazinediethanesulfonic acid.

(Received for publication, November 7, 1990)
Transcriptional Regulatory Region of the Human ICAM-1 Gene

sion levels of the endogenous ICAM-1 gene in a specific cell type and expression of ICAM-1-based reporter gene constructs when transfected into these cells thus demonstrate the biologic relevance of the identified transcriptional regulatory regions.

MATERIALS AND METHODS

Cells—HK were purchased from Clonetics (San Diego, CA) and grown without feeder cells in a serum-free, low calcium medium (KGM, Clonetics). Cells were used for experiments after the second passage, at 80–90% confluence. A431, a human epidermoid squamous cell carcinoma line (24) was obtained from American Type Culture Collection (Rockville, MD). B6 and B7 were derived from a tyrosine kinase deficient fibroblast cell line previously utilized in transfection studies (25). A431 and B6 cells were cultured in Dulbecco’s modified Eagle’s medium (4.5 g/liter glucose) (GIBCO) supplemented with 3 mM l-glutamine, 10% heat-inactivated fetal bovine serum, 100 units/ml penicillin, 0.25 mg/ml amphotericin B, and 100 μg/ml streptomycin (all from GIBCO). All cells were grown at 37°C in humidified 5% CO₂.

In some experiments, HK and A431 cells were exposed to recombinant human IFN-γ (Amgen Biologicals, Thousand Oaks, CA) at 250 units/ml for 24 h prior to preparation of RNA.

Genomic Library Screening—A commercially prepared human lymphocyte cDNA (Stratagene, La Jolla, CA) was screened with a human ICAM-1 cDNA (26). This cDNA contains the bases 46–1891 of the cDNA published by Staunton et al. (6), which will be used as a reference of cDNA nucleotide positions in this report. The cDNA insert was subcloned from the original vector aH3M into pGEM3Z (Promega, Madison, WI) as previously described (27). The genomic library was prepared by partial SauIIIa digestion of circulating lymphocyte DNA with subsequent cloning of size-fractionated genomic fragments into the BamHI sites of the substitution vector bacteriophage XADASH (Stratagene). A 315-nucleotide XbaI/HindIII fragment of the ICAM-1 cDNA was 32P-labeled by random primer extension (specific activity of 10⁹ cpm/μg (28) and used to screen duplicate filters of 9 × 10⁶ recombinant phage plaques (28), giving a 99% probability for the presence of one 15-kb insert with relevant genomic material (29). Hybridizing clones were plaque purified in subsequent screening rounds. Single clone phage DNA was prepared from lysates (30) using polyethylene glycol precipitation, sodium dodecyl sulfate/EPTA phage lysis, and DNA extraction/precipitation (28). The purified DNA was then used for subsequent analytical and cloning procedures.

Restriction Enzyme Mapping and Southern Blot Analysis—Purified phage DNA was digested with various combinations of restriction endonucleases, the fragments analyzed by horizontal agarose gel electrophoresis, and a restriction map was created. The digests were blotted onto nylon membrane (Schleicher & Schuell), and Southern analysis (31) was carried out with two 32P-labeled fragments of the human ICAM-1 cDNA. A 315-nucleotide 5' and a 1.5-kb 3' fragment were generated from the cDNA by simultaneous digestion with XbaI, which released the cDNA from its cloning site in pGEM3Z, and HindII, which cuts once within the cDNA. Hybridization was carried out for 20 h at 42°C in 50% formamide, 2.5 × Denhardt’s solution (1 × = 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 0.02% Ficoll), 5 × SSC (1 × = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 1% sodium dodecyl sulfate, and 100 μg/ml tRNA. Autoradiography, using XAR-2 film and intensifying screen at −70°C, was performed after increasingly stringent washes.

DNA Sequencing—For DNA sequence analysis, EcoRI fragments of the recombinant phage clone were subcloned into pGEM3Z in a shotgun library (30); selected inserts were sequenced directly from denatured double-stranded constructs. Sequencing was carried out by the dideoxyxenucleotide chain termination method of Sanger et al. (32) using the Sequenase 2.0 reaction kit (U. S. Biochemical). Primers utilized were either complementary to the T7 or SP6 sites flanking the multiple cloning site of the vector (Promega) and were commercially synthesized oligodeoxyribonucleotides complementary to experimentally established sequences. The sequence of pG6-2.05, the EcoRI subclone containing the 5’ most coding region of the ICAM-1 gene, was screened for promoter/enhancer consensus sequences using the sequence analysis program MacVector (IBI, New Haven, CT).

RNA Preparation—HK and A431 cells were harvested at 80–90% confluence by exposing them to 0.05% trypsin, 0.5 mM EDTA (GIBCO) for 5–10 min at 37°C followed by two washes with phosphate-buffered saline. Total cellular RNA was prepared by guanidium isothiocyanate lysis and CsCl gradient ultracentrifugation (28). Because the ICAM-1 transcript is absent or extremely low in untreated HK and A431 cells, respectively, and because it is relatively abundant in both cell lines that bind ICAM-1, IFN-γ (27), total cellular RNA was utilized in both primer extension reactions and S1 nuclease protection assays at 10 and 25 μg/reaction, respectively.

**Primer Extension Reaction**—Primer extension reactions for the determination of the extent of the 5'-untranslated region (5'-UTR) of ICAM-1 mRNA were performed as described (33), using synthetic oligodeoxyribonucleotides complementary to portions of the ICAM-1 mRNA. The 40-mer 5’ GCT GGG AGC CAT AGC GAG GCT GAG GAT GTC ACT CTG AGT A 3’ was designated P11; complementarity of its 5’ end begins 12 nucleotides downstream of the translation start site (AUG) of the ICAM-1 mRNA, and extends 28 nucleotides upstream into the 5'-untranslated region (UTR) of the mRNA. The 40-mer 5’ CTT GGA ACA ACA GGG CCA GGA GCA GGA GTG CCG GCC GCA G 3’, which is complementary to nucleotides fully within the coding region of the ICAM-1 mRNA located 67 nucleotides downstream of the AUG, was designated P21. The primers were 32P-end-labeled using T4 polynucleotide kinase (28) and hybridized with total cellular RNA from IFN-γ-treated or untreated A431 cells and HK. Incubation was for 2 h at 50°C in a solution containing 1 M NaCl, 0.5 M HEPE, pH 7.5, 1 mM EDTA, pH 8.0 RNA and primer were then coprecipitated and resuspended in 0.5× M KCl, 5 mM EDTA, 100 mM dithiothreitol, 4 mM dNTP, pH 8.3. 40 units of avian myeloblastosis virus-reverse transcriptase (Boehringer Mannheim, West Germany) were added to a final reaction volume of 25 μl and the primer extension reaction was carried out for 90 min at 42°C. The reaction was terminated with 1 μl of 0.5 M EDTA, pH 8.0, and reaction products were precipitated, resuspended in denaturing buffer, heat denatured, and fractionated by denaturing acrylamide gel electrophoresis. The size of the primer extension products was determined by counting the number of nucleotides in a concurrently run sequencing ladder. This sequencing reaction was performed on pG6-2.05 using the same 5' end primer. The 3' end primer was identical to the restriction enzyme primer allowing identification of specific bases which correspond to the cap site. After fractionation, the gel was dried and autoradiography performed as described above.

**S1 Nuclease Protection Assay**—A single-stranded 5' end-labeled DNA probe complementary to the ICAM-1 mRNA was generated from the genomic subclone pG6-2.05 as described (33). This 187-nucleotide single-stranded probe extends from 12 nucleotides downstream to 175 nucleotides upstream (3' end) of the translation start site of the ICAM-1 mRNA and was calculated to extend at least 100 nucleotides 5' of the candidate transcription start sites. The end-labeled probe was hybridized with total cellular RNA of either IFN-γ-treated or untreated A431 cells and HK for 20 h in a solution of 80% formamide, 0.4 M NaCl, 0.04 M PIPES, pH 6.4, 1 mM EDTA, pH 8.0. The hybridization mix was subsequently exposed to 300–1000 units/ml S1 nuclease for various times at 37°C in 0.58 M NaCl, 0.05 M sodium acetate, pH 4.5, 4.5 mM ZnSO₄, 20 μg/ml denatured salmon sperm DNA. The reaction was terminated by adding 1/4 volume 4 M ammonium acetate, 20 mM EDTA, pH 8.0, 40 μg/ml tRNA, and the reaction products precipitated, resuspended, and analyzed by denaturing acrylamide gel electrophoresis and autoradiography as described above.

**Production of Reporter Gene Constructs Using Various Portions of the ICAM-1 5'-Flanking Region and Chloramphenicol Acetyltransferase (CAT) Vectors**—CAT reporter gene constructs were designed and engineered to contain various portions of the ICAM-1 5'-flanking region contained within pG6-2.05 and were produced in vectors which either provided or lacked an SV40 transcription enhancer region (Fig. 4A). pHS-CAT-B was produced in the following manner. A 282-nucleotide HindIII/SstI fragment extending 277 base pairs 5' of the identified transcription start site (see below, "Results") was isolated. This fragment, designated HS, also contains by sequence analysis the putative functional TATA box, two potential Sp1 binding sites, and three potential CAAT boxes (see "Results" for nucleotide number locations). The promoterless CAT expression vector pCAT-Basic (Promega) was digested with HindIII/AccI, each cutting single sites (HindIII 5' of AccI) in the multiple cloning site located in the CAT coding region. The construct containing vector and HS were ligated, ensuring the same 5'-3' orientation for both the ICAM-1 HS fragment and the CAT coding region. The
insert from the phage arms and, additionally, cut it into nine fragments of 0.3, 0.4, 0.85, 1.2, 1.6, 1.7, 2.0, 3.8, and 5.4 kb representing 17.5 kb of genomic material. The assignment of the fragments was determined by additional digests with BamHI, SalI, and XbaI and various combinations of these enzymes. Southern blotting revealed that the 315-nucleotide Xbal/HindIII 5' fragment of the human ICAM-1 cDNA hybridized with both the 2.05- and the 5.4-kb EcoRI fragments of $\lambda$6, and 1.53-kb HindIII/XbaI 3' fragment of the human ICAM-1 cDNA hybridized only with the 3' most (0.85 kb) EcoRI fragment of $\lambda$6. The sequencing of exonic material of these fragments after cloning them into pGEM3Z, using T7/Sp6 promoter primers or oligodeoxynucleotides derived either from the ICAM-1 cDNA or from sequencing data, revealed mRNA coding stretches of 124, 263, and 203 nucleotides for the 2.05-, 5.4-, and 0.85-kb fragments, respectively. Within the exon sequences of pg6-2.05 are contained the first ATG codon with reference to the described cDNAs (6, 26) and 67 nucleotides of ICAM-1 protein coding region (Fig. 2). Furthermore, all the previously determined nucleotide sequence upstream from the translation start site, as documented in the published cDNAs, is found immediately 5' of the ATG codon. At the 3' end of the 0.85-kb fragment a consensus sequence for a splice donor site has been identified. The 263 nucleotides of coding region contained in pg6-5.4 are identical in sequence to cDNA nucleotides 125-388 and are flanked at their 5' and 3' ends by consensus sequences for splice acceptor and donor sites, respectively (data not shown), indicating that this region corresponds to a complete exon. The exon sequences in the insert of subclone pg6-0.85 are identical to cDNA nucleotides 389-592 and are bordered, at their 5' end, by a consensus splice acceptor site. The 3' end of this fragment has no associated splice donor site but contains a terminal SauIII site. However, since the 3' end of this 0.85-kb subclone corresponds precisely to the 3' end of the entire $\lambda$6 genomic fragment, and since the genomic

![Fig. 1. The structure of the human genomic clone $\lambda$6.](image)

![Fig. 2. The nucleotide sequence of pg6-2.05, containing the first exon and candidate promoter region.](image)
library from which λ6 was cloned was prepared by a partial SauIIIa digest of human genomic DNA, it is most likely that this SauIIIa site is the product of a preparative enzymatic restriction digest of the genomic DNA within this exon rather than the formal 3′ end of this exon. Thus, λ6 contains the first 533 nucleotides of ICAM-1 protein coding region organized over two complete exons and one partial exon, and all the 5′ UTR nucleotides of the ICAM-1 cDNA published to date.

Within λ6, the correspondence of exon regions to various protein domains of the final gene product is strikingly similar to that previously described for other members of the immunoglobulin superfamily (5). The 67 nucleotides of protein coding region contained within the first exon correspond to the first 22 and 1/3 codons for the 27-residue signal peptide of the ICAM-1 protein (6). The remainder of the signal peptide is encoded by the second exon, which also contains 82 and 2/3 of the 88 codons which determine the N-terminal immunoglobulin-like domain (6). Since pGA6-2.05 contains the 5′ most coding region that is represented in the cDNA, it was considered likely to contain candidate 5′-flanking transcriptional regulatory regions and was sequenced in its entirety (Fig. 2). Within the 2045 nucleotides of genomic material are contained the 67 nucleotides of protein coding region (see above), 1392 nucleotides of 5′-flanking region, and 586 nucleotides of the first intron. Consensus sequences for promoter elements, e.g. TATA and CAAT boxes and Sp1-binding sites, are found within the 5′-flanking region appropriately located with respect to the identified cap site (see below). Additionally, consensus sequences for a number of enhancer elements can be found in both the 5′-flanking and first intronal regions of pGA6-2.05. These include two potential interferon-stimulated response elements at nucleotides -1307 to -1294 and +465 to +478 (36); two potential glucocorticoid receptor elements at -1328 to -1314 and +658 to +642 (37); a potential NFκB site at -528 to -519 (38); two potential AP2 sites at -917 to -910 and -506 to -499 (39); and several potential AP1 sites (40).

Characterization of the 5′-UTR and Cap Site—Primer extension was performed using reverse transcriptase after end-labeled, complementary oligodeoxyribonucleotides had been hybridized to total cellular RNA of either untreated or IFN-γ-treated HK or A431 cells (Fig. 3A). No specific extension products were observed if total cellular RNA of untreated HK or A431 cells was used as template. The extension of PE1 produced specific bands of 51 and 52 nucleotides when reacted with RNA from IFN-γ-treated cells. Since the 5′ end of PE1 is complementary to the base located 12 nucleotides downstream of the translation start site AUG, these bands correspond to a 5′-UTR of 39 or 40 nucleotides in length. PE2, whose 5′ end is complementary to the base 67 nucleotides downstream of the AUG, produced bands of 106 and 107 nucleotides (Fig. 3B). Thus, extension studies utilizing non-overlapping primers confirm a 5′-UTR of 39–40 nucleotides for ICAM-1 transcripts generated under these experimental conditions, that is by HK and A431 cells after IFN-γ treatment.

To demonstrate that the full 5′-UTR was contained within the identified putative first exon, S1 nuclease protection assays were performed. A single-stranded probe of 187 nucleotides, complementary to the 5′ region of the ICAM-1 mRNA, was generated from the appropriate strand of pGA6-2.05. The 5′ end of this probe, like the primer PE1, is complementary to the base located 12 nucleotides downstream of the ICAM-1 AUG. 5′ end-radiolabeled probe was allowed to hybridize to total cellular RNA from IFN-γ-treated A431 cells, and S1 nuclease was subsequently added. S1 nuclease completely degraded single-stranded probe hybridized to irrelevant tRNA, but when hybridized with the A431 RNA, produced protected fragments or probe slightly greater than 50 nucleotides in length (Fig. 3C), consistent with a 5′-UTR of 39–40 nucleotides and in agreement with the primer extension data. Omission of S1 nuclease from the reaction resulted in the preservation of the full-length, single-stranded probe. In additional experiments using the same probe, protected fragments identical in size to those seen with A431 RNA were observed using RNA from IFN-γ-treated HK. RNA from untreated HK and untreated A431 cells did not protect the probe from S1 digestion (data not shown).

The primer extension and S1 nuclease protection experiments utilizing IFN-γ-induced transcripts identify a cap site within the genomic fragment of pGA6-2.05 that is 39–40 nucleotides 5′ of the ICAM-1 AUG (Fig. 2). Sequence analysis of the region immediately upstream from this cap site reveals a TATA box (41) within a GC-rich region at nucleotides -30

**FIG. 3.** Determination of the transcription initiation site of the ICAM-1 gene utilized in epithelial cells after stimulation with IFN-γ. A, autoradiography of a primer extension reaction with primer PE1. 32P-End-labeled PE1 was hybridized to total cellular RNA of IFN-γ-treated or untreated HK and A431 cells as well as to tRNA and then extended with reverse transcriptase as described under “Materials and Methods.” Specific extension products are marked by the arrow. B, autoradiography of a primer extension reaction with primer PE2. A sequencing reaction was run concurrently for the determination of the size of the extension products. The sequencing reaction was performed on pGA6-2.05 using a primer whose 5′ end coincides with that of PE2. Thus, the precise candidate nucleotides corresponding to the cap site for ICAM-1, as well as the length of the primer extension products, can be determined. Specific extension products are marked by the arrow. C, autoradiography of an S1 nuclease protection assay. A single-stranded, 5′-end-labeled DNA probe complementary to the ICAM-1 mRNA extending 12 nucleotides downstream (5′ end of probe) to 175 nucleotides upstream (3′ end of probe) of the translation start site of the ICAM-1 mRNA was hybridized to total cellular RNA of IFN-γ-treated A431 cells and to tRNA and then each either exposed (+) or not exposed (−) to S1 nuclease as described under “Materials and Methods.” Bands corresponding to specific protected probe fragments are marked by the arrow. Length of protected probe fragments was determined by concurrently electrophoresed sequencing reactions and radiolabeled nucleotide length markers.
to −25 (+1 designated as the guanine 40 nucleotides 5′ of the ICAM-1 AUG). Additionally, both consensus Sp1-binding sites (−196 to −191 and −59 to −54 (37)) and potentially utilized CAAT boxes (−262 to −254 and −154 to −146 (42) and −53 to −48 (43)) can be identified within the first 300 nucleotides of the 5′-flanking region identified by these studies.

The 5′-Flanking Region of the ICAM-1 Gene Contains a Potent Functional Promoter and an Upstream Region Which Confers Tissue Appropriate Repression of Transcription—Whether the candidate promoter region surrounding the identified transcription initiation site could drive expression of a heterologous reporter gene was first determined. A 282-nucleotide HindIII/SstI (HS) fragment of pGM6-2.05 (nucleotides −277 to +5 of Fig. 2), containing the identified cap site and candidate basal promoter elements described above, was inserted immediately upstream of the CAT reporter gene contained in the promoterless expression vector, pCAT-Basic (Fig. 4A). This inserted HS fragment functioned as a potent promoter of CAT expression when pHs-CAT-B was transiently transfected into both murine B6 fibroblasts and the squamous cell carcinoma line, A431 (Fig. 4B). In multiple transfection analyses, the relative expression of CAT enzyme using pHs-CAT-B was consistently as good or better than that observed with the positive control SV40 promoter/enhancer construct, pCAT-Control. The fact that expression of this hybrid construct was high in A431 cells, which normally express no or very low levels of ICAM-1 (27), raised the possibility that in engineering pHs-CAT-B, a potent ICAM-1 promoter had been divorced from flanking, constitutively utilized repressor elements. Therefore, CAT reporter constructs were produced using additional portions of the ICAM-1 5′-flanking region, and these constructs (Fig. 4A) were utilized in transient transfections to determine: 1) if sequences immediately 5′ of the HS fragment might contain regions that physiologically repress ICAM-1 gene transcription in the proper cellular context (construct pBS-CAT-P); and 2) if repression of CAT expression in A431 cells could be demonstrated with the inclusion of additional 5′-flanking regions, could this repression be overridden with the addition to the repressed construct of an SV40 enhancer (construct pBS-CAT-ENH). Furthermore, analysis of the ICAM-1 genomic region immediately 5′ of the HS fragment of pGM6-2.05 revealed not only consensus sequences for multiple potential cis-acting regulatory elements (see above), but also additional candidate cap sites with appropriately located basal promoter elements (potential cap site, nucleotides −284 to −278; potential TATA box, −313 to −307, Fig. 2). For this reason, and because previously defined ICAM-1 cDNA sequences indicated the existence and utilization of alternative transcription initiation sites in the expression of the ICAM-1 gene (see below, “Discussion”), reporter constructs were also devised to assess the ability of the region immediately flanking the HS fragment to function as a promoter, both without and with the assistance of an SV40 enhancer (Fig. 4A, constructs pBH-CAT-P and pBH-CAT-ENH, respectively). Results of CAT expression utilizing these additional constructs (Fig. 4C) in transient transfections into A431 cells reveal the following: 1) the 1165 nucleotide BS fragment, which includes both the identified potent promoter region contained in fragment HS and the immediately 5′ 883 nucleotides, confers markedly repressed expression of CAT (lane 3) when compared to expression driven by HS alone (lane 5); 2) while inclusion of an SV40 enhancer alone leads to no CAT expression (lane 6), addition of this enhancer to the repressed BS construct completely overrides the observed repression (lane 7); and 3) the 883 nucleotide upstream BH region displays no promoter function, either without (lane 4) or with (lane 8) the addition of the SV40 enhancer. These BH constructs show no promoter activity when transfected into murine B6 fibroblasts, as well (data not shown). These data demonstrate that while the ICAM-1 5′-flanking region contains a potent promoter region, it also contains physiologically relevant repressor elements that are both functional and capable of being overridden. Thus, in A431 cells, in which ICAM-1 gene expression is constitutively very low but which can be markedly upregulated

**Fig. 4.** A, schematics of pGM6-2.05, displaying specific restriction fragments of the 5′-flanking region of the ICAM-1 gene and ICAM-1-based CAT reporter gene constructs. The transcription start site identified in these studies is indicated by the bent arrow; the ICAM-1 and CAT protein coding regions are indicated by hatched and shaded boxes, respectively, and the restriction sites utilized, their relative locations, and resultant fragment sizes are noted. For details of hybrid plasmid construction, see "Materials and Methods." B, assay of function of the candidate ICAM-1 promoter region contained within the 282-nucleotide HS fragment. Murine B6 fibroblasts (lanes 1–3) and A431 cells (lanes 4–6) were transiently transfected with CAT-reporter gene constructs and assayed for CAT expression as described under "Materials and Methods." pCAT-Basic (lanes 1 and 4), the parent plasmid of pHs-CAT-B, has no endogenous promoter or enhancer. pCAT-Control (lanes 2 and 5) contains the SV40 promoter and enhancer regions, and thus serves as a positive control for CAT expression. CAT enzymatic activity after transfection with pHs-CAT is shown in lanes 3 and 6. The percentage of enzymatic conversion to acetylated forms of chloramphenicol are: pCAT-Basic: 0.4% (lane 1), 1.25% (lane 4); pCAT-Control: 18.8% (lane 2), 17.5% (lane 5); pHS-CAT: 26.0% (lane 3), 34.4% (lane 6). C, comparative assessment of promoter activity of various portions of the 5′-flanking region of the ICAM-1 gene after transfection into A431 cells. Constructs utilized were those generated by insertion of fragments BS, BH, or HS upstream of CAT reporter plasmids that either did not (−) or did (+) contain the SV40 enhancer (ENH), as schematically represented in A. Transfection, preparation of lysates, and assays of CAT expression were performed as described above. Expression of negative control pCAT-Basic and positive control pCAT-Control are shown in lanes 1 and 2, respectively. Expression of test constructs are as follows: pBS-CAT-P, lane 3; pBH-CAT-P, lane 4; pHS-CAT-B, lane 5; pCAT-Enhancer, a CAT expression vector containing only the CAT protein coding region and the SV40 enhancer (Promega), lane 6; pBS-CAT-C, lane 7; and pBH-CAT-C, lane 8.
Upon exposure to specific signals, concordance of the biology of endogenous ICAM-1 gene expression levels and transient transfection expression assays is achieved when larger, inherently more physiologic portions of the ICAM-1 gene 5'-flanking region are utilized in the hybrid expression constructs.

**DISCUSSION**

We have characterized the genomic structure of the human ICAM-1 5'-transcriptional regulatory region and have identified a potent, functional promoter within this region that is flanked by a functional repressor. Structural analysis has also revealed the exonal organization of that portion of the gene which encodes the N terminus of the ICAM-1 protein, specifically the signal peptide and the distal immunoglobulin-like domain, which contains the primary binding site for both lymphocyte function-associated antigen 1 and the major group of rhinovirus.

The protein coding region of the first exon contains 67 of 81 nucleotides (22 and 1/3 of 27 codons) that encode the cDNA-deduced signal peptide (6), and the second exon contains the remaining 14 nucleotides encoding the signal peptide and 247 of 264 nucleotides (82 and 2/3 of 88 codons) that encode the N-terminal immunoglobulin-like domain (6). Virtually identical exonal organization of signal peptide and distal immunoglobulin-like domains have been reported for other members of the immunoglobulin superfamily, e.g. for the variable (V) regions of immunoglobulin H, \(\kappa\), and \(\lambda\) chain genes (44), as well as for the V-regions of TCR \(\alpha\), \(\beta\), and \(\gamma\) chain genes (45). The three EcoRI fragments of \(\lambda\) that contain hybridizing, exonal material have also been detected in genomic Southern analysis (23). The exonal information of the ICAM-1 gene appears to be dispersed over a larger region of genomic DNA than predicted (6). Apart from \(\lambda\), we have also isolated another genomic recombinant phage clone that contains an additional 300 nucleotides of ICAM-1 protein coding sequence. Together, these two recombinant phage clones represent approximately 30 kb of ICAM-1 genomic sequence, but contain only 52.3% of the ICAM-1 protein coding region, indicating that the ICAM-1 gene is considerably larger than previously thought.

Utilizing RNA isolated from two epithelial cell types (HK and A431 cells) after IFN-\(\gamma\) treatment, we have identified a cap site that is appropriately located with respect to a consensus TATA box and predicts a 5'-UTR of 39-40 nucleotides. The 5'-end of the ICAM-1 cDNA isolated by Staunton et al. (6) extends an additional 17 nucleotides upstream from our experimentally determined cap site (5'-end of the Staunton et al. (6) cDNA at nucleotide -17, Fig. 2), and predicts a 5'-UTR of at least 57 nucleotides. Since this cDNA sequence was obtained from a phorbol myristate acetate-stimulated HL60 cDNA library, the inclusion of these additional 17 nucleotides in the 5'-UTR may have resulted from an mRNA generated using an alternate transcription start site located slightly upstream in the genomic sequence. Since this 17 nucleotide region is identical in sequence and location in both the cDNA and the \(\lambda\) subclone, it is highly unlikely that this discrepancy can be explained by alternative splicing or the existence of additional intron-separated ICAM-1 exonal regions further upstream. Indeed, probing of Northern blots which contain ICAM-1 transcripts with fragments of \(\lambda\) representing an additional 4.5 kb of upstream genomic sequences have failed to detect any hybridizing signals. From the genomic sequence reported here, and if the Staunton et al. (6) ICAM-1 cDNA represents a complete cDNA, such an alternative transcription initiation site would be located only 12 nucleotides downstream from the identified TATA box, with no alternative, appropriately distanced (25-30 nucleotides) TATA consensus sequence within the adjacent clone.

Repetitive S1 nuclease protection assays utilizing RNA isolated from IFN-\(\gamma\)-treated epithelial cells have failed to detect protected fragments which would correspond to a 5'-UTR of 57 nucleotides in length. However, in some primer extension studies, we have detected minor specific bands corresponding to a 5'-UTR of precisely this length (data not shown). Thus, while a predominant ICAM-1 transcript induced by IFN-\(\gamma\) in epithelial cells appears to utilize a transcription initiation site located 39-40 nucleotides 5' of the translation start site, there are indications that an alternate cap site which produces a transcript with a slightly longer 5'-UTR may also be present. However, additional upstream consensus cap sites with appropriately located promoter elements (contained within the BH fragment, Fig. 4) failed to demonstrate any functional promoter activity in transient expression assays of reporter gene plasmids, even with the addition of an SV40 enhancer. Nevertheless, these putative alternative transcription initiation sites (located 17 nucleotides apart) cannot explain why multiple ICAM-1 mRNA species ranging from 1.9 to 3.3 kb are usually detected in Northern blot analysis (6, 23, 26). These considerable mRNA length differences may reflect post-transcriptional modifications of the ICAM-1 mRNA. There is little indication, however, that alternative splicing affecting either the 5'-UTR or coding region could contribute to these multiple RNA species, since the isolated cDNAs have identical protein coding sequences and, at least with regard to the 5'-end of the mRNA, the cDNA sequences are completely contained within the characterized genomic DNA. Splicing events exclusively affecting the 3'-UTR of ICAM-1 mRNA have not been ruled out.

When inserted into a heterologous reporter gene construct, the immediate 5'-flanking region of the ICAM-1 gene demonstrates strong promoter activity when transfected into murine B6 fibroblasts and A431 cells. The chloramphenicol conversion rates for the ICAM-1 pHs-CAT-B construct are uniformly equivalent or better than those seen with the positive control SV40 promoter/enhancer CAT plasmid, pCAT-Control. Interestingly, this ICAM-1 promoter construct was also expressed at high levels in A431 cells, a cell line that shows no or only low constitutive expression of ICAM-1 under standard culture conditions (27). While these differences in expression might have been attributable to the absence of physiologic regulatory mechanisms in transient transfection expression systems, transfection experiments utilizing the additional ICAM-1-based reporter constructs pHs-CAT-P and pHs-CAT-C demonstrate that this discrepancy is explained by the dissociation of the basal ICAM-1 promoter from neighboring regulatory DNA sequences. These flanking sequences clearly contain cis-acting regions that, in the appropriate cellular context, physiologically repress ICAM-1 gene transcription. The absence of these presumptive repressor elements, as in the pHs-CAT-B construct, thus allows strong promoter activity even without the addition of normally required inducers of expression, such as IFN-\(\gamma\). Indeed, this observed constitutive repressor activity can be completely overcome with the addition of an SV40 enhancer to the appropriate reporter construct. Furthermore, preliminary experiments indicate that this constitutive repressor activity observed with the pHs-CAT-P construct can also be overridden by treatment of transiently transfected A431 cells with IFN-\(\gamma\), a signal which rapidly induces endogenous ICAM-1 gene expression in this and numerous other cell types. It is

K. Degitz, L. Lian-Jie, and S. W. Caughman, unpublished results.
therefore of interest that within the 5'-flanking region of the human ICAM-1 gene are contained potential interferon-responsive elements, glucocorticoid receptor-binding sites, an NFκB consensus element, and AP1 and AP2 sites, regions which may be involved in the regulated expression of this gene. The exact biologic roles played by these potential elements as well as other regions involved in the constitutive and tissue-specific regulation of ICAM-1 gene expression are currently under investigation.

Acknowledgments—We gratefully thank Mark Udey, John Stanley, and Stephen Katz for their helpful reviews of this manuscript and Ann Lemoine, Atlanta Veterans Affairs Medical Center, for her excellent work in its preparation.

REFERENCES
1. Dustin, M. L., Rothlein, R., Bhan, A. K., Dinarello, C. A., and Springer, T. A. (1986) J. Immunol. 137, 245–254
2. Rothlein, R., Dustin, M. L., Marlin, S. D., and Springer, T. A. (1986) J. Immunol. 137, 1270–1274
3. Bierer, B. E., and Burakoff, S. J. (1988) FASEB J. 21, 2584–2590
4. Makgoba, M. W., Sanders, M. E., and Shaw, S. (1989) Immunol. Today 10, 417–422
5. Williams, A. F., and Barclay, A. N. (1988) Annu. Rev. Immunol. 6, 381–405
6. Staunton, D. E., Marlin, S. D., Stratowa, C., Dustin, M. L., and Springer, T. A. (1988) Cell 52, 925–933
7. Makgoba, M. W., Sanders, M. E., Ginther, L. E., Dustin, M. L., Springer, T. A., Clark, E. A., Mannoni, P., and Shaw, S. (1988) Nature 331, 86–88
8. Marlin, S. D., and Springer, T. A. (1987) Cell 51(5), 813–819
9. Boyd, A. W., Wawryk, S. G., Burns, G. F., and Fecondo, J. V. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 5090–5099
10. Marlin, S. D., and Rothlein, R. (1989) J. Exp. Med. 167, 1323–1340
11. Faull, R. J., and Russ, G. R. (1989) Transplantation 48, 226–230
12. Marlin, S. D., Smith, J. A., and Struhl, K. (1989) J. Mol. Biol. 217, 847–861
13. Rodriguez, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarek, M. E., and McClelland, A. (1989) Cell 56, 839–847
14. Martindale, S. D., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D., and Springer, T. A. (1989) Cell 56(5), 849–853
15. Marlin, S. D., Staunton, D. E., Springer, T. A., Stratowa, C., Sommergruber, W., and Merluzzi, V. J. (1990) Nature 344, 70–72
16. Berendt, A. R., Simmons, D. L., Tansey, J., Newbold, C. I., and Marsh, K. (1989) Nature 341, 57–59
17. Sonn, J. P., Stade, B. G., Holzmann, B. Schwabl, W., and Riethmüller, G. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 641–644
18. Girard, D. J., Aaronson, S. A., Todaro, G. J., Stern, P., Kesey, J. H., Dosik, H., and Parks, W. P. (1973) J. Natl. Cancer Inst. 51, 1417–1423
19. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052
20. Caugham, M. W., and seed, B. (1988) Nature 331, 624–627
21. Caughman, S. W., Li, L. J., and degitz, K. (1990) J. Invest. Dermatol. 94, 228–235
22. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052
23. Caughman, S. W., and Seed, B. (1988) Nature 331, 624–627
24. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052
25. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052
26. Caughman, S. W., and Seed, B. (1988) Nature 331, 624–627
27. Caughman, S. W., Li, L. J., and Degitz, K. (1990) J. Invest. Dermatol. 94, 228–235
28. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052
29. Caughman, S. W., Hentze, M. W., Rouault, T. A., Harford, J. B., and Klauser, R. D. (1988) J. Biol. Chem. 263, 19048–19052