Endosialin, the antigen identified with monoclonal antibody FB5, is a highly restricted 165-kDa cell surface glycoprotein expressed by tumor blood vessel endothelium in a broad range of human cancers but not detected in blood vessels or other cell types in many normal tissues. Functional analysis of endosialin has been hampered by a lack of information about its molecular structure. In this study, we describe the purification and partial amino acid sequencing of endosialin, leading to the cloning of a full-length cDNA with an open reading frame of 2274 base pairs. The endosialin cDNA encodes a type I membrane protein of 757 amino acids with a predicted molecular mass of 80.9 kDa. The sequence matches with an expressed sequence tag of unknown function in public data bases, named TEM1, which was independently linked to tumor endothelium by serial analysis of gene expression profiling. Bioinformatic evaluation classifies endosialin as a C-type lectin-like protein, composed of a signal leader peptide, five globular extracellular domains (including a C-type lectin domain, one domain with similarity to the Sushi/ecto/ser pattern, and three EGF repeats), followed by a mucin-like region, a transmembrane segment, and a short cytoplasmic tail. Carbohydrate analysis shows that the endosialin core protein carries abundantly sialylated, O-linked oligosaccharides and is sensitive to O-sialoglycoprotein endopeptidase, placing it in the group of sialomucin-like molecules. The N-terminal 360 amino acids of endosialin show homology to thrombomodulin, a receptor involved in regulating blood coagulation, and to complement receptor C1qRp. This structural kinship may indicate a function for endosialin as a tumor endothelial receptor for as yet unknown ligands, a notion now amenable to molecular investigation.

The endosialin antigen was identified in a survey of normal and neoplastic human tissues conducted at the Ludwig Institute for Cancer Research in pursuit of new targets for antibody-based cancer therapies (37). The hallmark of monoclonal antibody (mAb) FB5, the probe used to discover endosialin (1), is its distinctive pattern of reactivity with human tissues. Thus, in a detailed study of biopsy and surgical specimens representing diverse cancer types, FB5 immunostaining was found primarily in tumor blood vessels and not in malignant tumor cells. Significantly, the antigen was not observed in all cancer samples examined, and even in cancers showing FB5-immunoreactive endothelial cells, the antigen was frequently detected with a heterogeneous pattern in the tumor vascular bed. Such a mixed pattern might be expected for a molecule involved in the reorganization of blood vessels in tissues such as cancers, in which areas of stable blood supply and histology are juxtaposed to regions of necrosis, hypoxia, excessive growth, tissue invasion, and remodeling. The normal tissues examined were unreactive with mAb FB5, including the blood vessel endothelium present in the respective organs.

The expression of the FB5 antigen by cultured normal and tumor cells was also investigated (1), revealing that the standard test cells for normal endothelial differentiation markers, cultured human umbilical vein endothelial cells, and microvascular endothelial cells derived from bone marrow and dermis are antigen-negative. Human umbilical vein endothelial cell cultures stimulated with a range of mediators known to induce the expression of endothelial activation antigens (2, 3) maintain their FB5-negative phenotype (1). Among a host of cultured epithelial, neuroectodermal, mesenchymal, and hematopoietic cell types tested, most were FB5 antigen-negative. Notable exceptions are short term cultures of normal fibroblasts and neuroblastoma cell lines, which consistently express the antigen in tissue culture (1), allowing the chromosomal assignment of an FB5 coding gene to human chromosome 11q13-qter, based on the serologic analysis of mouse-human neuroblastoma cell hybrids. The reason why cultured fibroblasts and neuroblastomas are FB5 positive in vitro, yet the corresponding cells in uncultured tissue sections are FB5 negative, is not known. Presumably, some of these findings reflect cell type-restricted, adaptive changes of the cell surface phenotype triggered by tissue culture factors (38); the mechanism underlying such induction and its relationship to FB5 antigen induction in tumor blood vessels are not known.

Methods are not generally available to examine the biochemical nature of antigens with restricted expression in tumor endothelium, because of the three properties of these cells: (i) they constitute a very minor fraction in most tumor tissues; (ii) they are not readily purified in sufficient numbers for direct
Identification of Endosialin

7409

protein analysis; and (iii) no cell culture model of tumor endothelium exists that faithfully reproduces its distinctive in vivo phenotype. With no source for antigen characterization in tumor endothelium, the target molecule for mAb FB5 was studied in cultured neuroblastoma and sarcoma cells and fibroblast cultures. In these cell types, the antigen is a unique 165-kDa glycoprotein, comprised of a core protein that migrates as a 95-kDa species on SDS gels and carries the mAb FB5-defined epitope, and abundant, highly sialylated O-linked carbohydrate moieties. There has been no proof that the same protein carries the FB5 epitope in tumor endothelium, but in recognition of the characteristic biochemical properties and the tumor endothelial expression pattern the molecule was designated endosialin.

In the absence of more detailed information about the molecular structure of endosialin, no clues were available regarding its potential function in cancer. Therefore, considering the keen interest in angiogenesis as a determinant of cancer progression and metastasis (4) and as a target for novel cancer therapies (5), this study aimed to clone the endosialin gene and provide the requisite probes and structural information to explore endosialin function in suitable model systems and in human cancers.

EXPERIMENTAL PROCEDURES

Cells and Antibodies—Cells line LA1-5s and MF-SH were grown in RPMI 1640 medium with 10% fetal calf serum and 20 IU/ml penicillin/ streptomycin; HeLa-S3 cells were grown in Nutrient mixture HAM F-12 with fetal calf serum and penicillin/streptomycin (Life Technologies, Inc.). The mAbs used were FB5 (mouse IgG2a) (1) and 9EG7 (ratIgG2a) against human β1-integrin chain (6).

Protein Purification and Edman Sequencing—Pelleted LA1-5s cells were lysed in RIPA buffer (150 mM NaCl, 1% Triton X-100, 0.1% SDS, 20 mM EDTA, 50 mM Tris/HCl, pH 7.4, 1 mM phenylmethylsulfonyl fluoride) at 55 °C, using 4 μl of total RNA were added to a mix of 4 μl of the display THERMO-RT buffer (Display Systems Biotech), 2 μl of a 5 mM dNTPmix, 7 μl of 3 M lof a 5 mM dNTPmix, 7 μl of 3 M lof 3 M MgCl2, 20 μl of reaction buffer alone (50 mM sodium phosphate, pH 7.0, 0.5% Triton X-100). Precipitates were washed, split into equal aliquots, and treated either with 4 μl of PNGase F (Flavobacterium meningosepticum, 10 milliunits sialidase from Arthrobacter ureafaciens, or a mixture of 10 milliunits sialidase and 0.5 milliunits O-glycosidase from Diplococcus pneumoniae (Roche Diagnostics) or with 25 μl of reaction buffer alone (50 mM sodium phosphate, pH 7.0, 0.5% Triton X-100). Precipitates were washed, eluted in Laemmli buffer and reduced with 50 μl dithiothreitol. The precipitated with mAbs or mIgG coupled to CNBr-Sepharose (Amersham Pharmacia Biotech) at a ratio of 2.8 mg/ml of gel. The precipitants were washed, split into equal aliquots, and treated either with 4 μl of PNGase F and analyzed by SDS-PAGE as described above. The products were resolved on a 2% (w/v) agarose gel and visualized by Coomassie staining. The endothelial bands were excised from the gel, incubated with trypsin solution (Promega, Madison, WI), and eluted peptides were subjected, after HPLC (Waters, Milford, MA) separation, to a 494 cLC ABI-PeptideElmer device. Peptides were desalted by ZipTip procedure and subjected to mass spectrometry analysis using a matrix-assisted laser desorption ionization-time of flight instrument (Voyager DE-STR, PerSeptive Biosystems/AB, Foster City, CA). Peptide samples were prepared using dihydroxybenzoic acid as matrix (7).

Northern Blot Analysis—The SV Total RNA Isolation System (Promega) was used for total RNA isolation according to the manufacturer’s instructions. Oligo(dT)-cellulose (Life Technologies, Inc.) was taken for polyA+ RNA isolation as described (8). For Northern blots, 4 μg of poly(A)+ RNA were electrophoresed on a 0.8% (w/v) agarose gel containing 20 mM MOPS, 5 mM sodium acetate, pH 6.6, and 1.11% formaldehyde. The RNA was blotted in 10× SSC (1.5 mM NaCl, 0.15 mM sodium citrate, pH 7.0) on a Hybond N+ membrane (Amersham Pharmacia Biotech) for 16–20 h and UV cross-linked. To detect endosialin mRNA, a 287-bp polymerase chain reaction fragment, comprising positions +2237 to +2524 of endosialin cDNA, was used.

RACE-Polymerase Chain Reaction—For reverse transcription 2 μg of total RNA were added to a mix of 4 μl of the display THERMO-RT buffer (Display Systems Biotech), 2 μl of a 5 mM dNTPmix, 7 μl of 3 M betain (Sigma), 1 μl of the display THERMO-RT Initiator Mix (total volume, 20 μl), and 1 μl of primer Est2243DIN2 (ACAGGTAAGCGTG- GCAGCAGCGGCTGC). For cDNA synthesis the mixture was incubated for 10 min at 65 °C, cooled to 42 °C, and incubated with 2 μl of the display THERMO-RT terminator solution for 40 min. The temperature was raised to 65 °C for 15 min, cooled to 37 °C, and then 1 μl of RNase Mix (‘5’ RACE System for Rapid Amplification of cDNA Ends Reagent Assay version 2.0, Life Technologies, Inc.) was added and incubated for 10 min at 37 °C for 30 min. The cDNA product was precipitated with 2.8 mg/ml of CNBr-Sepharose (Amersham Pharmacia Biotech) for 16–20 h and UV cross-linked. To detect endosialin mRNA, the cycling conditions described above. The products were resolved on a 2% (w/v) agarose gel and sequenced on an ABI PRISM 310 Genetic Analyser (PerkinElmer Life Sciences).

Transfection and Immunocytochemistry—Transient transfection of HeLa-S3 cells plated in chamber slides (Becton Dickinson, Heidelberg, Germany) was carried out with the transfection reagent FuGENE 6 and expression vector pMH (Roche Diagnostics), containing the complete endosialin coding sequence as empty vector. Cells for antigen expression 24 h after transfection, using immunocytochemistry with mAb FB5 or control IgG essentially as described (9). Antibody binding was detected with Alexa Fluor 488-conjugated goat anti-mouse IgG (Fab')2 fragment (Molecular Probes, Eugene, OR).

Immunoprecipitation Assays and Carbohydrate Analysis—Immunoprecipitations were performed as described (6). For glycosylation studies, cells were metabolically labeled with 400 μCi of [35S]methionine/400 μCi of [3H]cytochrome/cysteine-free minimal essential medium (Life Technologies, Inc.) for 16 h. Cell lysates were precleared with protein A-Sepharose (Amersham Pharmacia Biotech) and precipitated with mAbs or mIgG coupled to CNBr-Sepharose (Amersham Pharmacia Biotech) at a ratio of 2.8 mg/ml of gel. The precipitants were washed, split into equal aliquots, and treated either with 4 μl of PNGase F (Flavobacterium meningosepticum, 10 milliunits sialidase from Arthrobacter ureafaciens, or a mixture of 10 milliunits sialidase and 0.5 milliunits O-glycosidase from Diplococcus pneumoniae (Roche Diagnostics) or with 25 μl of reaction buffer alone (50 mM sodium phosphate, pH 7.0, 0.5% Triton X-100). Precipitates were washed, eluted in Laemmli buffer and reduced with 50 μl dithiothreitol. Signals were detected after SDS-PAGE by exposure to Kodak BIOMAX MR film. For O-sialoglycoprotein endopeptidase digestion, equal aliquots of LA1-5s cell detergent lysates were precipitated with mAb FB5 coupled to CNBr-Sepharose (2.8 mg/ml of gel) or with mAb 9EG7 against integrin β1-chain coupled to protein A-Sepharose (1 mg/ml of gel) via rabbit anti-rat IgG (1 mg/ml of gel) for 2 h at 4 °C. The precipitates were washed, split into equal aliquots, and treated either with 4 μl of pig microvascular endothelial endopeptidase from Pasteurella hemolytica (activity: 1 μl cleaves 10 μg of glycoprotein A/h at 37 °C; Cedarlane Laboratories, Hornby, Canada) or reaction buffer only (50 mM Tris/HCl, pH 7.4), for incubation for 1 h at 37 °C. Proteins were eluted and analyzed by SDS-PAGE as described above.

Database Mining and Sequence Analyses—Amino acid sequence information for endosialin peptides was used to perform iterative BLAST, FASTA, and Prosite pattern data base searches against public sequence data bases. The same algorithms, including gapped BLAST and PSI_BLAST (10), have been applied for further analysis of endosialin cDNA and amino acid sequence. The SEG program was applied for searching low complexity regions in protein sequences with the following search parameters: window length, 25; trigger complexity, 3.0; and extension complexity, 3.3 (11). Known functional sequence domains matching sequence segments in endosialin were examined in the PFAM and SMART (13) domain libraries.

RESULTS

Protein Purification of Human Endosialin and Identification of the cDNA—We used immunocytochemistry and enzymerelinked immunosorbent assays with mAb FB5 to confirm that human umbilical vein endothelial cell cultures and cultures of microvascular endothelial cells from dermis and bone marrow lack endosialin expression. Therefore, the FB5 antigen-positive neuroblastoma cell line LA1-5s was selected as a source for
Identification of Endosialin

Endosialin cDNA Encodes the FB5 Cell Surface Antigen—Bioinformatic analysis identifies endosialin as a typical precursor sequence for a eukaryotic cell surface protein, resulting evolutionarily from domain shuffling. The endosialin protein displays a considerable number of disparate functional segments (Fig. 5A). First, there is an N-terminal sequence of about 20 amino acids that is similar to a signal leader peptide for export to the endoplasmic reticulum and can be detected by the SIGNALP2 suite of Nielsen et al. (14). Next, there are two major segments, a more N-terminal globular portion with five distinct domains comprising amino acids 30–360, and a more C-terminal portion comprising amino acids 361–757 and classified almost exclusively as low complexity region by the SEG program (11). A more detailed analysis of domain structure by comparison with hidden Markov models in libraries of described sequence domains, such as PFAM (12) and SMART (13), as well as sequence similarity searches in data bases with the BLAST and PSI-BLAST tools (10) that rely on the concept of a common evolutionary ancestor among sequentially homologous sequences. By this approach, we found statistically significant hits for a C-type lectin domain (residues 29–157, E <

Endosialin Is a C-type Lectin-like Membrane Receptor—Fig. 2. cDNA and deduced amino acid sequence of human endosialin. The sequence of full-length cDNA was assembled from ESTs, derived based on the five tryptic peptide sequences derived from Edman analysis of purified endosialin (double-underlined amino acids) and from 5′-RACE analyses. The putative N-terminal signal sequence and the potential polyadenylation signal are underlined. The predicted transmembrane segment is boxed, and a single N-linked glycosylation site is marked by a circle.
Identification of Endosialin

A staining of endosialin-transfected HeLa-S3 cells (transfectants. Photomicrograph of indirect immunofluorescence dosialin expressing cell lines (LA1-5s and MF-SH; 4 m left GAPDH. Positions of size markers (in kilobases) are shown to the 271, E SMART; residues 316–350, Alexa 488-conjugated goat anti-mouse IgG. fication, 200 200 RNA/lane) was probed with 32P-labeled endosialin cDNA or a probe for GAPDH. Positions of size markers (in kilobases) are shown to the left.

10−5 with PFAM) and three EGF-like domains (residues 235–271, E < 10−5 with PFAM; residues 274–311, E < 10−2 with SMART; residues 316–350, E < 10−3 with PFAM). In accordance with SMART assignments, the latter two EGF repeats may be Ca2+ binding. The region 176–230 is probably a Sushi/SCR/CCP domain. Although the E value in a PFAM search is only 0.72, the typical motifs in the Sushi seed alignment with the endosialin segment are well conserved, including all necessary cysteines for disulfide bridges and a number of critical prolines, glycines, and a tryptophan.

A more detailed analysis of the C-terminal, low complexity portion of endosialin, covering amino acids 360–757, showed that the local amino acid composition is dominated by the preferential occurrence or absence of certain residue types. In accordance with analyses of known three-dimensional protein structures (15), there is little reason to expect such polypeptide segments to assume stable tertiary structures with sterically predefined binding or catalytic activity. It is not surprising that secondary structure tools (16, 17) predict almost exclusively coil preferences for this region. Segment 363–380 is highly negatively charged, with 6 glutamate and 7 aspartate residues, and it may serve a nonspecific role in facilitating a binding contact with a charged ligand molecule. Using the software tools DAS (18) and TOPPPRED2 (19), the existence of a single hydrophobic transmembrane region (amino acids 686–706) can be detected with strong significance; thus, endosialin appears to be a type I membrane protein. The extracellular region spanning amino acids 391–660 is much richer than the data base average (20) in prolines (21.9%), threonines (9.3%), and serines (9.3%), but it contains surprisingly few charged residues (DEKR total only 9.6%), glycines (2.2%), asparagines (1.1%), and phenylalanines (1.5%). There is no truly sequentially similar protein sequence in the data base. The small intracellular portion 707–757 is compositionally rich in proline (11.8%) but depleted of phenylalanine (none) and hydrophobic residues in general (LVIFM total only 17.6%). This sequence region is also unique compared with other known protein sequences.

Although the prediction of O-glycosylation sites is not very reliable even with the most advanced algorithms (21), we supplied the endosialin sequence to the NetOGlyc 2.0 server offered by J. E. Hansen. This program predicts as probable O-glycosylation sites Thr-60 (in the C-type lectin domain) and a series of 34 serine and threonine residues located in segment 400–669, which probably does not have a globular structure.

Endosialin Is Related to Thrombomodulin and Complement Receptor C1qRp—The N-terminal endosialin segment between amino acids 1 and 360 shares 39% sequence identity with the corresponding region in the precursor protein of human thrombomodulin (also referred to as fetomodulin, CD141, or emb:CAA29045.1) and 33% sequence identity with the human complement receptor C1qRp (also known as lymphocyte antigen 68, antigen AA4, or emb:CAC00597.1), respectively. The matching region encompasses the C-type lectin domain, the putative Sushi domain, and the three EGF-like repeats (Fig. 5B). Moreover, the positions of cysteine residues critical for disulfide bridges are highly conserved among the three proteins, consistent with a shared three-dimensional structure in this segment. Finally, the three proteins show a conserved WIGL consensus motif, a feature found in several additional cell surface proteins that modulate endocytosis (22).

Endosialin Is Conserved in Nonhuman Species and Maps to Human Chromosome 11q13—A number of mouse and rat ESTs, originating from embryonic tissue sources and comprising fragments of 186–521 bp were identified in public data base searches. These ESTs show 82–92% identity to the corresponding human endosialin cDNA sequence, thus suggesting the presence of a close homologue in nonhuman species. Further data base searches identified a human genomic clone (RP11–755F10, accession number AP000759) assigned to chromosome 11q13, which contains more than 75% of the endosialin coding sequence (positions +521 to +2533) in a putative single exon and in this portion is 99% identical with the endosialin cDNA. This chromosomal assignment for the endosialin confirms and refines the assignment to 11q13-ter made for the FB5 antigen using somatic cell hybrids (1).

Carbohydrate Analysis of Recombinant Endosialin—In several human cell types, endosialin shows a high degree of O-glycosylation and sialic acid content (1), and we have examined the glycosylation pattern of recombinant endosialin in transiently transfected HeLa-S3 cells (Fig. 6A). When tested by immunoprecipitation with mAb FB5 immobilized to CNBr-Sepharose, a 165-kDa protein is detected in these cells, but not in mock-transfected HeLa-S3 cells. Pretreatment of matrix-bound antigen with O-glycosidase or sialidase led to a marked reduction in the size of the endosialin band on SDS gels, with a pattern identical to endogenous endosialin in LA1-5s cells (Fig. 6A). Thus, sialidase treatment alone yields a 120-kDa protein species, and combined treatment with sialidase and O-glycanase generates the putative 95-kDa core protein. No shift in the mobility of endosialin on SDS gels was observed.

Fig. 3. Northern blot analysis of endosialin. RNA from two endosialin expressing cell lines (LA1-5s and MF-SH; 4 mg of poly(A)+RNA/lane) was probed with 32P-labeled endosialin cDNA or a probe for GAPDH. Positions of size markers (in kilobases) are shown to the left.

Fig. 4. Cell surface expression of endosialin on HeLa-S3 cell transfectants. Photomicrograph of indirect immunofluorescence staining of endosialin-transfected HeLa-S3 cells (A; magnification, 200×) or vector control-transfected, parental HeLa-S3 cells (B; magnification, 200×), respectively, with mAb FB5. Binding was detected with Alexa 488-conjugated goat anti-mouse IgG.
Identification of Endosialin

**FIG. 5.** Domain architecture of endosialin and sequence comparison to human thrombomodulin and complement receptor C1qRp. A, domain architecture of endosialin. The diagram shows the arrangement of the N-terminal signal leader peptide (triangle), the C-type lectin domain (C-LEC), the domain similar to a Sushi/SCR/CCP domain (S), three EGF-like repeats (EGF), and the transmembrane region (TM), flanked by two segments with low complexity regions, namely a sialomucin-like sequence (MUCIN) and a short, putative cytoplasmic tail (CYT). B, protein sequence alignment of endosialin with thrombomodulin (36) (SwissProt P07204) and the complement receptor C1qRp (22) (TREMBL U94333). The alignment shows the N-terminal 360 amino acids of endosialin (end marked by black slash), which include the C-type lectin domain, the Sushi-like domain, and the three EGF-like repeats. In this region, endosialin shares 39% identity with the matched region of thrombomodulin and 33% identity with C1qRp (white letters on blue background), respectively. Nearly all cysteine residues in the five domains are conserved among the three proteins (red). The WIGL consensus motif is depicted in green.

After PNGase F treatment, consistent with the notion that minor, if any, N-linked sugar moieties are present at the single Asn-Xaa-(Ser/Thr) consensus site at position Asn-628. It is noteworthy that endogenous endosialin in LA1-5s cells does not show the marked heterogeneity in glycosylation, revealed by broad bands on SDS gels, that is typical for a number of other glycoproteins with abundant O-glycosylation. The appearance of two closely spaced endosialin bands in the HeLa-S3 transfectants may be indicative of limited heterogeneity of O-glycosylation in this particular cell type, because the two bands merge into a single band after deglycosylation.

*Mucin-like Endopeptidase Susceptibility—*Because O-sialoglycoprotein endopeptidase (OSGE) has been shown to cleave specifically highly O-glycosylated mucin-like molecules (23, 24), we examined whether endosialin is susceptible to this enzyme. To this end, detergent extracts of metabolically labeled LA1-5s cells were purified with mAb FB5 and treated with OSGE for 1 h at 37 °C prior to SDS-PAGE. We found that OSGE completely degraded endosialin (Fig. 6B), whereas a control glycoprotein, β3-integrin, was largely insensitive to OSGE under these conditions.

**DISCUSSION**

We have cloned the gene coding for the human endosialin core protein based on several lines of evidence. First, transfection of endosialin cDNA into HeLa-S3 cells leads to induction of cell surface reactivity with the cognate mAb FB5. Secondly, mAb FB5 detects the expected 165-kDa sialoglycoprotein in the transfectants, with a core protein migrating as a 95-kDa species on SDS gels. Third, the endosialin gene maps to chromosome 11q13, refining the previous assignment to 11q13-qter based on serologic analysis of somatic cell hybrids with mAb FB5. Finally, the amino acid sequence provides the requisite attachment sites for abundant O-glycosylation of endosialin.

The present study is not the only evidence linking endosialin induction to vascular endothelial cells in human cancer. Rather, an independent line of investigation, aimed at dissecting comprehensive gene expression profiles in cancers with the serial analysis of gene expression (SAGE) method, has implicated endosialin-specific ESTs in tumor angiogenesis. St. Croix et al. (25) employed SAGE to survey about 32,500 unique gene transcripts for differential expression in purified tumor endothelium versus endothelium of normal tissue. The study identified 46 transcripts elevated at least 10-fold in tumor endothelium, and the highest and most consistent up-regulation was noted for a hitherto uncharacterized gene, deposited as ESTs of unknown function in data bases and designated tumor endothelial marker 1 (TEM1) by St. Croix et al. (25). We show here that TEM1 is the endosialin gene.

In their study, St. Croix et al. (25) confirmed the SAGE data and showed by *in situ* RNA hybridization that TEM1/endosialin expression in vivo is selective for tumor endothelium, with no detection in the malignant epithelial cells of colorectal and other cancer, a range of normal organs, and a panel of cultured tumor cell lines. Nevertheless, TEM1/endosialin mRNA was not unique to tumor blood vessels because endothelial cells associated with wound healing and corpus luteum formation also showed gene expression by *in situ* RNA hybridization.

Taking together the endosialin expression data generated...
with mAb FB5 (1) and the RNA expression data derived from SAGE and in situ RNA hybridization studies (25), a consistent picture of endosialin induction in tumor endothelium emerges. Such a close agreement between epitope and protein/mRNA expression is instructive for a molecule like endosialin, which to display a heterogeneous pattern of expression as a rule rather than exception. Considering the marked heterogeneity in FB5 staining in tumor vessels (1), it may be surprising that the report on TEM1/endosialin mRNA in situ hybridization (25) does not address this aspect. One explanation would suggest that endosialin mRNA is uniformly expressed in tumor endothelium and that only subsets of endothelial cells accumulate the endosialin protein or protein with the FB5 epitope; however, as outlined above, there is currently no evidence supporting such a mechanism. Alternatively, it may be argued that mRNA in situ hybridization is the more sensitive detection method for endosialin and that FB5 immunostaining marks tumor endothelium with the highest levels of protein/epitope density rather than a distinct tumor endothelial cell type. Finally, there may simply be differences in the design of the two available studies on endosialin expression.

Our analysis shows that the N-terminal extracellular portion of endosialin is likely to consist of five domains with globular structure, and in this region homology to thrombomodulin and the complement receptor C1qRp, with 39 and 33% amino acid sequence identity, respectively, is observed. Importantly, a pattern of cysteine residues is highly conserved among the three proteins and provides a scaffold of anchoring points for disulfide bridges, presumably imposing similar three-dimensional structures on each protein. Thrombomodulin exhibits a C-type lectin domain, six EGF-like repeats, and a serine/threonine-rich region in the extracellular domain. Unlike endosialin, thrombomodulin is expressed on a variety of normal cell types and on normal vascular endothelium serves as a receptor for thrombin, modulating the coagulation cascade and triggering the thrombin-activatable fibrinolysis inhibitor (30). The thrombin interaction depends on the EGF repeats 5 and 6 of thrombomodulin (31, 32), which are not present in endosialin, making it unlikely that the two molecules overlap in this particular function. The complement receptor C1qRp is also expressed on a wider variety of cell types than endosialin, including macrophages, neutrophil granulocytes, and normal vascular endothelium. The extracellular portion of C1qRp is composed of a C-type lectin domain, a tandem of five EGF-like repeats, and a mucin-like region. The molecule binds the complement factor C1q, mannos-binding lectin, and pulmonary surfactant protein A (22). On myeloid cells, C1qRp mediates phagocytosis of antibody-coated particles, but the function of C1qRp on normal endothelial cells is largely unknown (33).

Predictions about the membrane-proximal extracellular domain focus on fewer structural features. Thus, the high content of proline, threonine, and serine residues indicates likely attachment points for more than 30 O-linked carbohydrate side chains consistent with the sialomucin-like features of endosialin (28, 34). Certain sialomucins have been implicated in sequestering growth factors to the plasma membrane and presenting these factors to endothelial cell surface receptors (34,
Identification of Endosialin

35); endosialin may be a candidate for such a function unique to tumor endothelium.

Apart from further biochemical investigations that will benefit from the availability of mAb FB5 and recombinant endosialin protein, study of defined alterations in endosialin gene expression and structure are now possible, including genetic manipulations in mice and rats, which harbor a close homologue of endosialin. In conclusion, key experiments are now within reach to explore the endosialin system in human physiology and disease and to extend these investigations into experimentally more amenable rodent models. A better understanding of endosialin may even enable new approaches to cancer detection and treatment.

Acknowledgments—The expert technical assistance of Claudia Eiberle, Elfriede Müller, Werner Rust, and Kai Zuckschwerdt is gratefully acknowledged. Prof. Junichi Sakamoto kindly provided cell line MF-SH. Dr. Elisabeth Stockert kindly supplied purified mAb FB5. We fully acknowledged Prof. Klaus Pfizenmaier throughout the work. Thank Dr. Lloyd J. Old for valuable contributions and suggestions. Sven

REFERENCES

1. Rettig, W. J., Garin-Chesa, P., Healey, J. H., Su, S. L., Jaffe, E. A., and Old, L. J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 10832–10836
2. Lasky, L. A. (1992) Science 258, 964–969
3. Carles, T. M., and Harlan, J. M. (1994) Blood 84, 2068–2101
4. Hanahan, D., and Folkman, J. (1996) Cell 86, 353–364
5. Hanahan, D., and Weinberg, R. A. (2000) Cell 100, 57–70
6. Lenter, M., Uhlig, H., Hamann, A., Jeno, P., Imhof, B., and Vestweber, D. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 9651–9655
7. Hatai, E., Brett, S., Zoepfah, A., Ashman, K., Tentsch, U., Ahorn, H., Murphy, C., Schweigerer, L., and Fotsis, T. (2000) Adv. Exp. Med. Biol. 476, 239–248
8. Weller, A., Isenmann, S., and Vestweber, D. (1992) J. Biol. Chem. 267, 15176–15183
9. Rettig, W. J., Garin-Chesa, P., Beresford, H. R., Oetgen, H. F., Melamed, M. R., and Old, L. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3110–3114
10. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
11. Wootten, J. C. (1994) Comput. Chem. 18, 269–285
12. Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L., and Sonnhammer, E. L. (2000) Nucleic Acids Res. 28, 263–266
13. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P., and Bork, P. (2000) Nucleic Acids Res. 28, 231–234
14. Nielsen, H., Brunak, S., and von Heijne, G. (1999) Protein Eng. 12, 3–9
15. Saqi, M. (1995) Protein Eng. 8, 1069–1073
16. Eisenhaber, F., Imperiale, F., Argos, P., and Frommel, C. (1996) Proteins 23, 157–168
17. Frishman, D., and Argos, P. (1997) Proteins 27, 329–335
18. Cserzo, M., Wallin, E., Simon, I., von Heijne, G., and Elofsson, A. (1997) Protein Eng. 10, 673–676
19. von Heijne, G. (1992) J. Mol. Biol. 225, 487–494
20. Brendel, V., Bucher, P., Nourbakhsh, I. R., Blaisdell, B. E., and Karlin, S. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 2003–2006
21. Hansen, J. E., Lund, O., Tolstrup, N., Goeley, A. A., Williams, K. L., and Brunak, S. (1998) Glycoconjug. J. 15, 115–130
22. Nepomuceno, R. R., Henschen-Edman, A. H., Burgess, W. H., and Tenner, A. J. (1997) Immunity 6, 119–129
23. Abdullah, K. M., Udh, E. A., Sheewen, P. E., and Mellors, A. (1992) Infect. Immun. 60, 56–62
24. Sutherland, D. R., Abdullah, K. M., Cyopick, P., and Mellors, A. (1992) J. Immunol. 148, 1458–1464
25. St. Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K. E., Montgomery, E., Lal, A., Riggins, G. J., Lengauer, C., Vogelestein, B., and Kinzler, K. W. (2000) Science 289, 1197–1202
26. Hakomori, S. (1991) Curr. Opin. Immunol. 3, 646–653
27. Girling, A., Bartikova, J., Burchell, J., Gendler, S., Gillett, C., and Taylor-Papadimitriou, J. (1989) Int. J. Cancer 43, 1072–1076
28. Morgan, S. M., Samulowitz, U., Darley, L., Simmons, D. L., and Vestweber, D. (1999) Blood 93, 165–175
29. Puri, K. D., Finger, E. B., Gaudernack, G., and Springer, T. A. (1995) J. Cell Biol. 131, 261–270
30. Kokame, K., Zheng, X., and Sadler, J. E. (1998) J. Biol. Chem. 273, 12135–12139
31. Wood, M. J., Sampoli Benitez, B. A., and Komives, E. A. (2000) Nat. Struct. Biol. 7, 200–204
32. Wang, W., Nagashima, M., Schneider, M., Morser, J., and Neshime, M. (2000) J. Biol. Chem. 275, 22942–22947
33. Hess, K. L., Tudor, K. S., Johnson, J. D., Osati-Ashitian, F., Askew, D. S., and Cook-Mills, J. M. (1997) Exp. Cell Res. 236, 404–411
34. Sassetti, C., Van Zante, A., and Rosens, S. D. (2000) J. Biol. Chem. 275, 9001–9010
35. Goretzki, L., Burg, M. A., Grako, K. A., and Stallcup, W. B. (1999) J. Biol. Chem. 274, 16831–16837
36. Suzuki, K., Kusumoto, H., Deyashiki, Y., Nishio, J., Maruyama, I., Dash, K., Kawahara, S., Honda, G., Yamamoto, S., and Horiguchi, S. (1987) EMBO J. 6, 1891–1897
37. Old, L. J. (1996) Sci. Am. 275, 102–109
38. Rettig, W. J., and Old, L. J. (1989) Annu. Rev. Immunol. 7, 481–511