Characterization of SMOC-1, a Novel Modular Calcium-binding Protein in Basement Membranes*

Received for publication, April 19, 2002, and in revised form, July 7, 2002
Published, JBC Papers in Press, July 18, 2002, DOI 10.1074/jbc.M203830200

Christian Vannahme‡‡, Neil Smyth‡, Nicolai Miosge¶, Silke Göslingly, Christian Frie‡,
Mats Paulsson‡, Patrik Maurer¶¶, and Ursula Hartmann§§*

From the ‡Institute for Biochemistry, Medical Faculty, University of Cologne, Joseph-Stelzmann Strasse 52,
D-50931 Cologne and ¶Center for Anatomy, Department of Histology, University of Göttingen, Kreuzbergring 36,
D-37075 Göttingen, Germany

We have isolated the novel gene SMOC-1 that encodes a secreted modular protein containing an EF-hand calcium-binding domain homologous to that in BM-40. It further consists of two thyroglobulin-like domains, a follistatin-like domain and a novel domain. Recombinant protein in human cells showed that SMOC-1 is a glycoprotein with a calcium-dependent conformation. Results from Northern blots, reverse transcriptase-PCR, and immunoblots revealed a widespread expression in many tissues. Immunofluorescence studies with an antiserum directed against recombinant human SMOC-1 demonstrated a basement membrane localization of the protein and additionally its presence in other extracellular matrices. Immunogold electron microscopy confirmed the localization of SMOC-1 within basement membranes in kidney and skeletal muscle as well as its expression in the zona pellucida surrounding the oocyte.

BM-40 (also known as SPARC or osteonectin) was originally isolated from bone (1) but was subsequently found in a variety of other tissues (for review see Refs. 2 and 3). Its presence in basement membranes such as Reichert’s membrane or the basement membrane-rich Engelbreth-Holm-Swarm tumor together with the size of 40 kDa led to the name BM-40 (4).

BM-40 is a modular protein composed of three independently folded domains. The N-terminal domain contains about 50 amino acids of which 18 are negatively charged. The second module is homologous to follistatin (FS) with 10 cysteines in a typical pattern. The C-terminal extracellular calcium-binding (EC) domain has two EF-hand calcium-binding motifs, each with a bound calcium ion in the x-ray structure (5, 6). The EC domain of BM-40 additionally contains a binding site for several fibrillar collagen and the basement membrane collagen IV (7). PDGF also interacts with the EC domain of BM-40 (8) but in a calcium-independent manner, whereas collagen binding is calcium-dependent. Furthermore, the binding to PDGF is not influenced by the mutations that abolish collagen binding. The interaction of PDGF with BM-40 prevents binding to its receptor and thus growth factor signaling (9). Vascular endothelial growth factor was recently also shown to bind to BM-40 suggesting a general role in growth factor binding modulation (10).

BM-40 was reported to participate in the regulation of cell-matrix interactions, in particular influencing bone mineralization, wound repair, and angiogenesis. In vitro BM-40 inhibits cell adhesion, cell spreading, and cellular proliferation and regulates the expression of proteins involved in matrix turnover (reviewed in Ref. 2). It is highly expressed in some malignant tumors and was reported to play a crucial role in the tumorigenicity of human melanomas (11, 12).

Deletion of the gene coding for BM-40 in mice does not alter embryogenesis. However, postnatally BM-40-deficient mice develop cataracts with an abnormal cell surface-basement membrane interface (13–15) and severe osteopenia (16). These mice also display defects in wound healing (17, 18) and in tissue remodeling (19, 20).

Four homologous proteins have since been characterized, SC1/hevin (21), QR1 (22), tsc36/Flik (23), and testican-1 (24). SC1 is a glycoprotein present in synaptic junctions (21). Its expression in adult rat brain is up-regulated during reactive astrocytosis induced by mechanical trauma (25). The human ortholog (hevin), isolated from high endothelial venules (26), like BM-40, inhibits in vitro cell attachment and spreading of endothelial cells (27). Mice deficient in SC1 display no obvious phenotype (28). QR1 is expressed exclusively in the developing neuroretinal cells of quails and chickens where it becomes localized to the pericellular matrix (22). So far no homolog of QR1 has been found in mammals. It is most similar to SC1/hevin, but the N-terminal domain is not conserved, and expression patterns are completely different. tsc36/Flik is a transforming growth factor β1-induced protein originally identified in osteoblasts and a glial cell line (23, 29). Antisense treatment of chick embryos resulted in defects in axial patterning and holoprosencephaly presumably caused by attenuation of dorsalizing and neural-inducing signals during gastrulation (30) and suggesting a function similar to follistatin. The proteoglycan testican-1 was first isolated from human seminal fluid (31) and contains FS and EC domains. As with BM-40, the
EC domain of testican-1 binds calcium; however, no collagen binding activity could be detected (32). Two glycosaminoglycan attachment sites are present in the C-terminal domain (24). In the adult mouse, it is prevalent at post synaptic areas (33) and is developmentally regulated in the nervous system where it correlates with neuronal migration, axonal growth, and synaptogenesis (34). Similarly, in muscle development testican-1 becomes clustered at the neuromuscular junctions during post synaptic differentiation (35). In man, testican-1 is expressed by neurons and by endothelial cells of a variety of tissues (36). An unstable 130-kDa form is present in blood and converted to a smaller stable form by unidentified plasma serine proteases (37).

We recently identified three novel members of the BM-40 family, two of which share the domain organization of testican-1 and were termed testican-2 (38) and -3 (39). The third protein has a unique domain organization and was termed SMOC-1, where SMOC stands for secreted modular calcium-binding protein. The present work concerns its structure and expression.

MATERIALS AND METHODS

Isolation of cDNA Clones—A human fetal brain cDNA library (CLONTECH) was screened with two 187- and 239-bp 32P-labeled StyI fragments isolated from the human synovial membrane EST done 167131 (ATCC). Hybridization was carried out at 65 °C in aqueous solution (0.5 M sodium phosphate, pH 7.2, 7% SDS, 1 mM EDTA, 10% bovine serum albumin (BSA), and 0.2 mg/ml salmon sperm DNA). Filters were washed under low stringency conditions (40 mM sodium phosphate, pH 7.2, and 1% SDS) once at 55 °C for 15 min and twice at 65 °C for 15 min. Positive plaques were excised and rescreened. Three plaques from the final rescreeen were excised in vivo, yielding cDNA in the pDR2 vector. Plasmids were sequenced on both strands with flanking and internal primers using an ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit, and the products were resolved on an ABI Prism 377 Automated Sequencer (PE Biosystems). Analysis of the nucleotide sequence and homology searches in the dbEST (40) and GenBank™ data base were performed with the programs of the GCG package.

Protein Expression and Purification of Human SMOC-1—Primers were designed to amplify cDNA fragments spanning the full-length cDNA of human SMOC-1. Forward primer 5′-GCCCCTGATGCCACCCGCACCAAGCAG-3′ introduced an NheI restriction site at the 5′ side, and the reverse primer 5′-CAATGACTCTCGAGCTAGACGAGGCGTCCTTC-3′ introduced a stop codon together with an NcoI restriction site. PCR-amplified cDNA was cloned in the pCRII vector (Invitrogen) and sequenced by cycle sequencing. NheI/NcoI restriction fragments of the pCRII-SMOC-1 plasmid purified were cloned in the eukaryotic expression vector pCEP-Pu-His-Myc (41). Correct insertion of the construct in the pCEP-Pu-SMOC-1 was verified by sequencing. Plasmids were transfected into the human fibrosarcoma cells HT-1080 using an electroporator (Bio-Rad) according to the instructions of the manufacturer.

Growth and selection of transfected cells were carried out as described (32). Conditioned serum-free media of HT-1080 cells were collected and passed over a column of Talon Matrix (CLONTECH). Proteins were eluted in a linear gradient of 0–50% phosphate, pH 7.5, and the reverse primer 5′-CAAATGACTCTCGAGCTAGACGAGGCGTCCTTC-3′ introduced a stop codon together with an NcoI restriction site. PCR-amplified cDNA was cloned in the pCRII vector (Invitrogen) and sequenced by cycle sequencing. NheI/NcoI restriction fragments of the pCRII-SMOC-1 plasmid purified were cloned in the eukaryotic expression vector pCEP-Pu-His-Myc (41). Correct insertion of the construct in the pCEP-Pu-SMOC-1 was verified by sequencing. Plasmids were transfected into the human fibrosarcoma cells HT-1080 using an electroporator (Bio-Rad) according to the instructions of the manufacturer. Growth and selection of transfected cells were carried out as described (32). Conditioned serum-free media of HT-1080 cells were collected and passed over a column of Talon Matrix (CLONTECH). Proteins were eluted in a linear gradient of 0–250 mM imidazole in 50 mM sodium phosphate, pH 7.4, containing 300 mM sodium chloride. The yield of SMOC-1 was of the order of 500 μg/liter culture supernatant.

Circular Dichroism Spectroscopy—Circular dichroism spectra were recorded in a Jasco model 715 circular dichroism spectropolarimeter at 25 °C in thermostatted quartz cells of optical pathlength 1 mm. The molar ellipticity [θ] (expressed in degrees·cm2·dcm−1) was calculated on the basis of a mean residue molecular mass of 110 Da. The Ca2+ dependence of the circular dichroism spectrum was measured by addition of 2 mM CaCl2. Reversibility of the conformational change was tested by subsequent addition of 4 mM EDTA. A base line with buffer (5 mM Tris-HCl, pH 7.4) was recorded separately and subtracted from each spectrum.

Northern Blotting—Total RNA was extracted from various tissues of adult mice by the guanidinium-phenol method (42). Poly(A)+ RNA was prepared using Oligotex mRNA Midi Kit (Qiagen). 5-μg aliquots of poly(A)+ RNA from each tissue were electrophoresed on a 1% denaturing agarose gel and transferred onto Hybond-XL (Amersham Bio-

A 546-bp probe was generated by Nhel/DraIII-digestion of a PCR-amplified full-length construct using the forward primer 5′-GGCCCTGAGCCACCCGCACCAAGCAG-3′ and the reverse primer 5′-CAATGACTCTCGAGCTAGACGAGGCGTCCTTC-3′. This DNA fragment was labeled with [α-32P]dCTP following the manufacturer’s instructions (Takara). Hybridization was performed in 6× SSPE, 5× SSPE (750 mM sodium chloride, 50 mM sodium dihydrogenphosphate, 5 mM EDTA), 5× Denhardt’s (0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% BSA), and 1.5% SDS at 42 °C overnight. The blot was washed twice in 0.5× SSC (75 mM sodium chloride, 7.5 mM sodium citrate, pH 7.4), 0.1% SDS at 65 °C for 10 min and exposed overnight to a phosphorimager photoplate (Amersham Biosciences). Hybridization was carried out overnight in 1 × sodium chloride, 50 mM Tris-HCl, pH 7.5, 1% SDS at 65 °C with a PCR-derived 391-bp probe that was labeled with [α-32P]dCTP as described. The blot was washed twice with 0.1× SSC (15 mM sodium chloride, 1.5 mM sodium citrate, pH 7.4) for 20 min at 65 °C and exposed to an x-ray film (New RX, Fuji) for 10–15 min.

Production of Antiserum against SMOC-1—The rabbit was immunized with purified recombinant SMOC-1 (Pineda Antikörper Service). The antiserum did not react with any other known member of the BM-40 family in dot blots of native protein bound to nitrocellulose or in SDS-PAGE followed by immunoblot (results not shown).

SDS-PAGE and Immunoblot of Cell Culture Supernatants—Serum-free cell culture supernatants were collected, and 1-ml aliquots were precipitated by the addition of Triton X-100 to a final concentration of 0.5% trichloroacetic acid in water. The pellets were dissolved in SDS-containing sample mixture and electrophoresed under reducing conditions in a 12% SDS-polyacrylamide gel. After electrophoretic transfer to nitrocellulose, SMOC-1 was detected with the antiserum to SMOC-1 followed by a secondary antibody coupled to horseradish peroxidase (Dako). The enzyme product was detected by chemiluminescence.

Analysis of Glycosylation and Attachment of Glycosaminoglycans—For the removal of N-linked carbohydrates, recombinant SMOC-1 (1 μg) was digested in 50 mM Tris-HCl, pH 7.4, with 1 μl of PNGase F (200 milliunits/sample; Roche Molecular Biochemicals) at 37 °C for 2 h. A control sample was treated the same way without adding enzyme. Both samples were electrophoresed on a 12% SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue.

Equal aliquots (1 ml) of serum-free cell culture supernatants were precipitated by the addition of 9 ml of ethanol. Pellets were dissolved in TBS, 1% SDS and incubated at 95 °C for 15 min. After dilution to 0.1% SDS with TBS, Nonidet P-40 was added up to 0.5%, and the sample was incubated at 37 °C overnight with 1 μl of PNGase F (200 milliunits/sample). A control sample was treated the same way without adding enzyme. Both samples were separated on a 4–10% gel under reducing conditions and subjected to immunoblotting using the antiserum against SMOC-1.

To identify glycosaminoglycan chains, 250 ng of affinity purified SMOC-1 were incubated with either of the enzymes described above. Three μl of the respective reaction buffer was added in a 10-fold concentration, and digestions were performed at 37 °C overnight. Digestions
tions with chondroitinase ABC (20 milliunits/sample; Seikagaku) were performed in 20 mM Tris-HCl, pH 8, 50 mM NaCl, and 50 mM sodium acetate. Incubation with heparinase I (200 milliunits/sample; Sigma) and heparinase III (40 milliunits/sample; Sigma) was carried out in 20 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 4 mM CaCl₂. TBS was used for incubation with keratanase II (18 milliunits/sample; Sigma). Untreated and digested samples were separated on a 4–15% SDS-polyacrylamide gel and subjected to immunoblotting using an antibody against the Myc epitope of the tag (Santa Cruz Biotechnology).

The assessment of O-linked glycosylation was done using the digoxigenin glycan detection kit (Roche Molecular Biochemicals). 250 ng of affinity-purified SMOC-1 was treated with PNGase F as described above. Samples were separated on a 4–15% SDS-polyacrylamide gel and transferred to nitrocellulose. Further treatment was done according to the manufacturer’s instructions.

Immunohistochemistry on Mouse Tissue Sections —Freshly prepared mouse tissues were fixed for 1 h in 4% paraformaldehyde/PBS and after dehydration were embedded in paraffin. Microtome sections of 10 μm were cut, and the paraffin was removed. Epitopes were unmasked with bovine testes hyaluronidase (Sigma type IV-S, 300 μg/ml in PBS, pH 7.5) for 1 h. Immunostaining was conducted using a mouse monoclonal anti-SMOC-1 antibody and the appropriate secondary antibodies. The sections were counterstained with hematoxylin.

FIG. 1. Schematic overview of the in vivo excised clones (A) and complete cDNA sequence and deduced amino acid composition of human SMOC-1 (B). A, clone I obtained the complete sequence (3669 bp), and clones II and III covered only parts of the cDNA. The asterisk in A and B indicates the stop codon. The sequence is available from EMBL/GenBank®/DDBJ under accession no. AJ249900.
5.2, 30 min, 37 °C) and with proteinase K (5 μg/ml in 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 8 min, 37 °C) and again briefly fixed with 4% paraformaldehyde/PBS. After blocking with 1% BSA/TBS the sections were incubated with the rabbit antiserum to SMOC-1 followed by Cy-3-labeled anti-rabbit IgG (Dako) and observed under a Zeiss Axios-phot fluorescence microscope.

**Immunogold Electron Microscopy of Mouse Tissues**—Pieces of 1-mm³ size from the renal cortex of mouse kidneys, soleus muscle, and ovary were fixed in 4% paraformaldehyde and 0.5% glutaraldehyde for 15 min, dehydrated in a graded series of ethanol up to 70%, and embedded in the acrylic resin LR-Gold (London Resin Company). For electron microscopy, ultrathin sections were cut with a Reichert ultramicrotome and collected on Formvar-coated nickel grids. The procedure has been described previously in detail (43).

24-nm gold particles were prepared and directly coupled to the antibody against SMOC-1 according to procedures described previously (43). Tissue sections on nickel grids were incubated for 15 min at room temperature with PBS. Thereafter, the grids were incubated with gold-labeled antibodies against SMOC-1 diluted 1:50 with PBS for 1 h. The sections were rinsed with PBS, stained with uranyl acetate for 15 min, and lead citrate for 5 min and examined with a Zeiss (Leo EM906E) electron microscope. To exclude unspecific binding of the colloidal gold probes to anionic binding sites in tissue structures, control sections were incubated with the pure gold solution or with gold-coupled goat anti-rabbit antibodies (Medac) under the same conditions as described above. All controls were negative.

**RESULTS**

**Cloning and Characterization of Human SMOC-1 cDNA**—Analysis of sequence homologies between expressed sequence tags (EST) and human BM-40/osteonectin/SPARC revealed a set of four different clones with a 37% identity at the amino acid level. The homologous region encompassed 55 amino acids which included an EF-hand calcium-binding motif. The presence of two cysteines in the helices of the EF-hand is characteristic for EF-hands of the BM-40 family (6). To investigate whether this may represent a novel member of the BM-40 protein family, we used a cDNA fragment of an EST derived from human synovial membrane to probe a human fetal brain cDNA library. We could isolate three different clones (Fig. 1A). Clone I contained an insert of 3669 bp with an open reading frame of 1302 bp, a 5'-untranslated region (UTR) of 2113 bp, and a 3'-UTR of 2113 bp. The insert of clone II (1143 bp) covered nucleotides 1034–2176 of the full-length cDNA. The insert of clone III apparently had a mixed origin with only nucleotides 1172–1366 being identical to clone A. The 5'-UTR has a high GC content and lacks a TATA box. At the end of the 3'-UTR (residues 3648–3653) a consensus polyadenylation signal (AATAAA) is present. The isolated cDNA encodes a putative protein sequence of 434 amino acids with a calculated molecular mass of 48.2 kDa (Fig. 1B). Because of its modular composition and the presence of a calcium-binding domain (see below), we propose to call this novel gene product SMOC-1 for secreted modular calcium-binding protein-1.

**Domain Organization of SMOC-1**—A stretch of 26 amino acids at the N terminus of human SMOC-1 conforms well with the signal peptide consensus and ends with a signal peptidase cleavage site (44). SMOC-1 has no transmembrane-spanning hydrophobic region and is secreted from transfected cells. Mature SMOC-1 consists of 408 amino acids. Scrutiny of the sequence and comparison with other proteins allows the distinction of five domains, a follistatin-like FS domain, a thyroglobulin-like TY domain, a domain unique to SMOC-1, a EC domain and a second TY domain (Fig. 2). Residues 27–89 are homologous to the canonical FS domain. FS is composed of two subdomains with the second being similar to the Kazal domain (45). The first six cysteines of SMOC-1 can be aligned to cysteine residues 4–10 of the FS domain of BM-40 but also to the elastase inhibitor that represents an example of a “nonclassical” Kazal domain (46) (Fig. 3). The sequence similarity of the N-terminal domain of SMOC-1 is somewhat higher to the FS domain than to the Kazal domain, although the FS domain has an extension absent in both SMOC-1 and Kazal domains.

C-terminal to the FS domain, two TY domains (residues 90–160 and 226–293) are separated by 65 amino acids without homology to any known protein. The high content of aromatic amino acids implies that this region forms a folded domain with a hydrophobic core. A potential N-glycosylation site is present at Asn-214. The TY domains contain six cysteine residues including a characteristic CWCV tetrapeptide sequence (47). The structure of the TY domain of p41, the invariant chain of the major histocompatibility class II complex, was recently solved (48). The TY domain was defined by a first submodule with a short α-helix-β-strand arrangement and a second submodule with three strands forming a short antiparallel β-sheet. Both subdomains are stabilized by internal disulfide bonds. The structure-based alignment shows that all six cysteine residues and the features of secondary structure are conserved in both TY domains of SMOC-1 (Fig. 3). The C-terminal domain of SMOC-1 is homologous to the EC domain of BM-40, the characteristic amphipathic α-helix and the helix-loop-helix motifs of two EF-hands being conserved. However, in contrast to the EC domain of BM-40, which contains a variant EF-hand in addition to a canonical one, both EF-hands of SMOC-1 are canonical. Thus SMOC-1 is a modular protein built from several domains appearing in a number of other extracellular proteins, but also with a novel domain and a unique domain arrangement (Fig. 2).

**Structure of the SMOC-1 Gene**—The structure of the human SMOC-1 gene was elucidated by analysis of two HTGS clones (AL137789 and AL135747) both originating from the sequencing of chromosome 14. The SMOC-1 gene was mapped to 14q24.1. The gene spans about 150 kb from the translation start signal to the end of the known cDNA sequence. The coding region of the SMOC-1 gene consists of 12 exons (Fig. 4 and Table 1). Each domain of SMOC-1 is encoded by one or more exons, and the domain borders coincide with splice sites.

**Recombinant Expression of SMOC-1**—A SMOC-1 expression vector was constructed based on the pCEP-Pu plasmid (32) modified to code for an N-terminal hexahistidine peptide followed by a Myc tag and a cleavage site for factor X (pCEP-Pu-
His-Myc-FX (41)). The His-Myc-tagged SMOC-1 was expressed in the human fibrosarcoma cell line HT-1080 and could be purified from the cell culture supernatant by immobilized metal ion affinity chromatography on a cobalt column. Recombinant SMOC-1 migrated as a single band at 67 kDa in nonreducing SDS-PAGE (Fig. 5A). After reduction SDS-PAGE gave an apparent molecular mass of 75 kDa for His-Myc-tagged SMOC-1. The theoretical mass for His-Myc SMOC-1, calculated from the amino acid sequence, is 50,713.1 Da, and the discrepancy with the estimated mass indicates the presence of post-translational modifications and/or an abnormal behavior of SMOC-1 upon SDS-PAGE.

Therefore, we performed enzymatic digestions with the purified material using glycosidase PNGase F to remove N-linked carbohydrates and different glycosaminoglycan-lyases for removal of potential glycosaminoglycan chains. Digestion with PNGase F resulted in a shift in electrophoretic mobility corresponding to a slight loss of mass (Fig. 5B). Carbohydrates could still be detected on samples after incubation with PNGase F, indicating the additional attachment of O-linked sugar chains (results not shown). Treatment with chondroitinase ABC or metal ion affinity chromatography on a cobalt column.

Fig. 3. Structure-based alignment of SMOC-1 and related proteins. The sequence was aligned to the sequence of the FS and EC domains of BM-40, to the Kazal domain of the elastase inhibitor (46), and to the TY domain of p41 (48). Secondary structural elements as revealed by the structures of BM-40 (45) and p41 (48) are given below the consensus with β-sheets, α-helices, and EF-hands marked. Numbers above the sequence indicate the predicted disulfide bonds based on the disulfide linkage in BM-40 and thyroglobulin. Identical or conserved amino acids are shown on a black or shaded background, respectively. Calcium-coordinating residues are marked by arrows.

Fig. 4. Exon-intron structure of the human SMOC-1 gene. All exons are numbered and span the full coding region of the SMOC-1 gene. The open boxes for exon 1 and 12 indicate that the exact borders of the gene could not be determined. SP indicates the signal peptide.

His-Myc-FX (41)). The His-Myc-tagged SMOC-1 was expressed in the human fibrosarcoma cell line HT-1080 and could be purified from the cell culture supernatant by immobilized
Characterization of SMOC-1

Intron sequences in lowercase letters; exon sequences are in capital letters and divided into codon triplets. Amino acids at exon-intron borders are indicated with the position number based on the protein sequence deduced from the human cDNA.

### TABLE I

| Exon  | Domain          | Exon size | Splice donor | Intron size | Splice acceptor | Codon phase | Amino acid |
|-------|-----------------|-----------|--------------|-------------|-----------------|-------------|------------|
| 1     | 5'-UTR/SP       | 352       | CCCAGG gtaagt | 723         | acctag TTA TTA | 0           | Arg-33/Phe-34 |
| 2     | FS              | 166       | GC AAA G     | 11          | tttcag AT GCT G | I           | Asp-89     |
| 3     | TY              | 113       | ACC CAG gtagg | 222         | gacaag GTG CAG | 0           | Gln-126/Val-127 |
| 4     | TY              | 100       | GT TCA G     | 21          | cccag GT TCA G | I           | Gly-160    |
| 5     | NEW             | 48        | GG AAA G     | 14.4        | attcag AT GAC G | I           | Asp-176    |
| 6     | NEW             | 57        | GA GAT G     | 1.9         | cccctag AA ATC A | I           | Gly-195    |
| 7     | NEW             | 81        | AT TCA G     | 16.3        | ttcag AG AAA A | I           | Glu-222    |
| 8     | TY              | 193       | C ACA CG gtaagc | 0.5        | tccag C TAC GT | II          | Arg-286    |
| 9     | TY              | 83        | TA CCA G     | 1.8         | ttttag GC TGT C | I           | Gly-314    |
| 10    | EC              | 106       | T GGG AG     | 9.7         | gcagcag G TIC TC | II          | Arg-349    |
| 11    | EC              | 245       | AA GAA G     | 6.8         | cagtag GA GCC C | I           | Gly-429    |
| 12    | EC/3'-UTR       | 2122      |              |             |                 |             |            |

Conformation and Calcium-binding of SMOC-1—Circular dichroism spectroscopy of the tagged SMOC-1 was performed both in the presence and absence of calcium (Fig. 6). The spectra showed a distinct folding with low (5–6%) proportions of α-helix and predominant (~40%) β-structure. Calcium-induced a conformational change reminiscent of that seen for BM-40, indicating that the conserved EC domain is functional.

Tissue Distribution of SMOC-1 mRNA—Northern blot analysis of mRNA from different tissues of adult mice indicates a broad expression of SMOC-1. The strongest signals were seen in ovary, but signals were also detected in brain, thymus, heart, skeletal muscle, liver, and lung (Fig. 7). We confirmed these results by RT-PCR amplification of mRNAs combined with autoradiographic detection after Southern blotting (Fig. 8). A signal in testis could also be clearly detected. Spleen appears to be devoid of SMOC-1 transcripts.

Expression of SMOC-1 in Cultured Cells—A rabbit antiserum against the recombinant SMOC-1 was produced and tested for its reactivity with SMOC-1 and all other known members of the BM-40 family (BM-40, SC1, TSC36, testican-1, testican-2, and testican-3) in slot blots of native proteins bound directly to nitrocellulose as well as in Western blots obtained after SDS-PAGE of reduced proteins (results not shown). In all assays the antiserum reacted strongly with SMOC-1 and not at all with the other structurally related antigens.

The expression of SMOC-1 was tested in immunoblots of equal aliquots of conditioned media from cultures of a large panel of cell lines (Fig. 9). SMOC-1 was expressed by cells both of epithelial and mesenchymal origin. With most cell lines the major reactivity was seen as one or two bands migrating in the range of 70–90 kDa. Like recombinant SMOC-1, the native protein bears N-linked oligosaccharides because both bands show a shift to lower molecular weight upon digestion with PNGase F (Fig. 10). Whereas the lower band corresponds to the size of the recombinant protein, the upper band probably represents an isoform with an additional, unidentified post-translational modification.

Distribution of SMOC-1 within Mouse Tissues—The distribution of SMOC-1 in mouse tissues was investigated by indirect immunofluorescence on paraffin sections and compared with that of laminin as detected by an antiserum raised against mouse laminin-1 (Fig. 11). In skin both SMOC-1 and laminin were seen in the dermal-epidermal basement membrane zone and around capillaries. In addition antibodies against SMOC-1 showed a broader staining of the dermis and subcutis. In kidney SMOC-1 was detected associated with tubular and glomerular basement membranes. SMOC-1 staining surrounded skeletal muscle fibers and was detected in the inner meninges and capillaries and around defined neuronal cell populations in the brain. Taken together the results demonstrated a localization of SMOC-1 in or around basement membranes, but also in some additional tissue compartments that do not stain for the basement membrane marker laminin.

A particularly interesting staining pattern was seen in the ovary, where a strong signal was present at the periphery of the oocytes (Fig. 12). This was confirmed by immunoblotting of extracts from isolated ova (results not shown). No signal was seen in the basement membrane surrounding the follicle, indicating that SMOC-1 is not constitutively expressed in all basement membranes.

Association of SMOC-1 with Basement Membranes and Zona Pellucida—The localization of SMOC-1 to basement membrane zones was further analyzed by ultrastructural immunogold histochemistry (Fig. 13). This demonstrated that SMOC-1 is indeed a true basement membrane component. It is present in all basement membranes in the mouse kidney, i.e. of proximal and distal tubules, collecting ducts, and in the glomerular basement membranes and in Bowman’s capsule. Here SMOC-1 was found over the entire width of the basement membrane, including the lamina lucida, densa, and fibroreticularis. In the soleus muscle SMOC-1 is a component of the basement membranes of the myocyte as well as of the capillaries. In the mouse ovary, SMOC-1 is exclusively localized in the zona pellucida. Staining was seen in an area adjacent to the microvilli of the oocyte which extend into the zona pellucida.

DISCUSSION

SMOC-1 is a member of the BM-40 family of proteins as defined by containing a pair of a follistatin-like (FS) domain and an extracellular calcium-binding (EC) domain. The modular structure of this family is expanded in testicans and tsc36 where additional domains such as the thyroglobulin-like (TY) domain or a domain with partial similarity to van Willebrand factor type C domains have been inserted during evolution. The domain organization of SMOC-1 shows a further variation of this modular setup: SMOC-1 contains one FS, one EC, two TY domains, and a novel domain without known homologs. In all other members of the BM-40 protein family the FS domain is immediately followed by the EC domain, and both domains interact via a small surface (45). Although the interaction of the FS domain with the EC domain influences calcium binding
to the latter (49), the EC domain is functional and assumes the same structure when expressed separately (6). In SMOC-1 the FS and the EC domain are separated by the two TY domains which are themselves split by the novel domain.

Analysis of the gene structure shows that an intron is present at each domain border in SMOC-1. Domains that have become mobile during evolution are characterized by introns of the same phase at their domain borders (50), ensuring that the reading frame is maintained when the domain is inserted into an intron of another gene. Introns at the domain borders of FS, EC, and TY modules of the BM-40 family, including SMOC-1, are of phase I. However, positions and phases of introns that are located within the coding region of the domains are not conserved.²

From the amino acid sequence the EC domain of SMOC-1 can be predicted to be functional for calcium binding. The characteristic acidic residues at positions 1, 3, 5, 9, and 12 and the signatures for helices encompassing the calcium-binding loops are fully conserved for both EF hands. This prediction was confirmed experimentally with a conformational change observed for SMOC-1 when circular dichroism spectra were recorded in the presence and absence of calcium. Calcium binding is presumably important for the structure of SMOC-1 as seen for BM-40 which is in the calcium-bound form when present in the extracellular environment (49).

From the sequence of SMOC-1 three N-glycosylation sites can be predicted (Asn-153, Asn-214, and Asn-374). Digestion with PNGase F showed that one or more of these are used when SMOC-1 is recombinantly expressed in HT-1080 cells. The detection of carbohydrates in samples after PNGase F digestion indicates additional attachment of O-linked sugar chains. In

² P. Maurer, unpublished observations.
contrast to testican-1, -2, and -3, further members of the BM-40 family, no potential glycosaminoglycan attachment sites are present in the sequence. Accordingly, incubation of recombinant SMOC-1 with either chondroitinase ABC or heparinase I and III did not alter the migration behavior of the protein. However, a minor shift was seen after digestion with keratanase, indicating the presence of keratan sulfate. The demonstrated post-translational modifications may together account for the discrepancy between the calculated and the apparent molecular weight of the protein, but we cannot exclude an abnormal migration behavior of the protein. For the endogenous protein two bands could be detected in extracts of HT-1080 cells that both showed a smaller size after digestion with PNGase F. Because the lower band corresponds in size to the full-length recombinant protein, the second band is presumably not due to proteolytic processing but reflects additional post-translational modification.

Prediction of the biological function of the FS, TY, and EC domains in SMOC-1 is highly speculative and needs experimental verification. The FS domain is a widespread module not only found in follistatin and members of the BM-40 family but also present in the follistatin-related gene, in the complement proteins C6, C7, and factor I, in agrin, and in the transmembrane receptors tomoregulin-1/TMEFF1 and TMEFF2 (51–54). Whereas both follistatin and BM-40 bind growth factors, neither uses the FS domain for this (8, 55). The crystal structure of the FS domain revealed two subdomains, an N-terminal one with similarity to epidermal growth factor and a C-terminal subdomain that is homologous to the Kazal-type protease in-
hibitor domain (45). Based on the alignments it is not possible
to predict whether the FS domain of SMOC-1 has protease-
inhibiting activity, particularly as the sequence of the active
loop between Cys-8 and Cys-16 of the elastase inhibitor, the
most closely related Kazal domain, is not conserved in
SMOC-1.

Eleven copies of the TY domain are present in thyroglobulin,
but the function of these domains is unknown. TY modules
have spread into additional modular proteins. A subgroup of
TY-containing proteins including equistatin, the cysteine pro-
tease inhibitor ECI, saxiphin, and the major histocompatibil-
ity class II-associated invariant chain p41 (CD74) have proven
protease inhibitory function (56). However, nidogens and the
tumor-associated antigens GA-733-1 and -2 have no such ac-
tivity. Whereas testican-3 is able to inhibit membrane-type 1
matrix metalloproteinase this appears not to be due to the TY
domain (57). Future studies will ascertain the importance of
these domains in SMOC-1.

SMOC-1 mRNA is found in a wide variety of tissues, and the
protein is secreted by established cell lines of both epithelial
and mesenchymal origin. The broad tissue distribution was
confirmed also in immunofluorescence microscopy where
SMOC-1 was often found associated with basement membrane
structures. Basement membranes are mainly formed by a col-
lagen IV network in which an independent laminin complex is
intermingled (for review see Ref. 58). Nidogens and perlecan
link the two networks. Many matrix proteins are located in
basement membrane zones at the light microscopic level, al-
though only a few are integral components at the ultrastruc-
tural level (59). SMOC-1 was, however, found in all layers of
the basement membrane, and it lies within the basement mem-
brane, rather than being associated with its surface, indicating
that it is a true basement membrane component. Laminin-1,
nidogen-1, and collagen type IV are similarly localized over the
entire width of basement membranes (59). In contrast to these,
SMOC-1 is not a ubiquitous basement membrane component
because some basement membranes, like the one surrounding
the ovarian follicle, are devoid of SMOC-1.

BM-40 binds to collagen IV and fibrillar collagens through its
EC domain (5, 60). This binding interaction is relevant for the
localization of BM-40 in vivo as type I collagen-deficient mice
do not retain BM-40 in the extracellular matrix (61). The col-
lagen-binding epitope on the EC domain was mapped to five
crucial residues located on the opposite site of the EF-hands on
the N-terminal \( /H9251 \) helix and the loop that connects the EF-
hands. The x-ray structure of a BM-40 variant with increased
collagen affinity revealed that a flat surface forms the binding
epitope, and its diameter matches that of a collagen triple helix
(7). Although it is tempting to speculate that binding of
SMOC-1 to collagen IV underlies its basement membrane lo-
calization, the residues used for collagen binding in BM-40 are
not conserved in SMOC-1. In particular the linker region be-
tween the EF-hands is longer and contains six lysines (Fig. 3).
Binding of SMOC-1 to collagen IV in a similar manner as BM-40 is thus questionable. SMOC-1 is also found in tissue compartments lacking laminin. In ovaries SMOC-1 is localized to the zona pellucida, an extracellular matrix surrounding the oocyte. This coat physically separates oocyte and granulosa cells (62, 63) and is not only crucial for the survival of the oocyte but also for successful fertilization and the passage of early embryos through the oviduct. In mice it is composed of three glycoproteins (ZP1, ZP2, and ZP3) that make up at least 95% of the total zona protein (64). Whereas mice lacking ZP1, ZP2, or ZP3 exhibit abnormal folliculogenesis and varying degrees of infertility (65–67), it remains to be seen if SMOC-1 plays a role in the biology of the zona pellucida and the fertilization process.

Acknowledgment—We are grateful for the expert technical assistance of Johanna Hülsmann.

REFERENCES

1. Termine, J. D., Kleinman, H. K., Whitson, S. W., Conn, K. M., McGarvey, M. L., and Martin, G. R. (1983) Coll. 26, 89–105.
2. Lane, T. F., and Sage, E. H. (1994) FASEB J. 8, 163–173.
3. Maurer, P. (1996) in Guidebook to the Calcium-binding Proteins (Celo, M. R., ed.) pp. 169–171, Oxford University Press, Oxford.
4. Dziadek, M., Paulsson, M., Aumailley, M., and Timpl, R. (1986) Eur. J. Biochem. 161, 455–464.
5. Maurer, P., Hohenadl, C., Hohenester, E., Göhring, W., Timrl, R., and Engel, J. (1995) J. Mol. Biol. 253, 347–357.
6. Hohenester, E., Maurer, P., Hohenadl, C., Timrl, R., Janssion, N. J., and Engel, J. (1996) Nat. Struct. Biol. 3, 67–73.
7. Sasaki, T., Hohenester, E., Göhring, W., and Timrl, R. (1998) EMBO J. 17, 1625–1633.
8. Göhring, W., Sasaki, T., Heldin, Ch.-H., and Timrl, R. (1998) Eur. J. Biochem. 255, 69–66.
9. Raine, W. E., Lane, T. F., Iruela-Arispe, M. L., Ross, R., and Sage, H. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1288–1285.
10. Kupprian, C., Motamed, K., and Sage, H. (1998) J. Biol. Chem. 273, 29635–29640.
11. Ledda, M. F., Adris, S., Bravo, A. I., Kairiyama, C., Bover, L., Chernajovsky, Y., Mordoh, J., and Podhajzer, O. L. (1997) Nat. Med. 3, 171–176.
12. Nischt, R., Wallich, M., Rebetanz, M., Baumann, P., Krieg, T., and Mauch, C. (2001) Cancer Let. 162, 223–230.
13. Gilmour, D. T., Lyon, G. J., Carlton, M. B., Sanes, J. R., Cunningham, J. M., Anderson, J. R., Hogan, B. L., Evans, M. J., and Colledge, W. H. (1998) EMBO J. 17, 1860–1870.
14. Norsøe, K., Clark, J. D., Syed, N. A., Basu, A., Heber-Katz, E., Sage, E. H., and Timpl, R. (1999) Invest. Ophthal. Vis. Sci. 39, 2674–2680.
15. Norsøe, K., Lane, W. K., Clark, J. D., Sage, E. H., and Howe, C. C. (1998) Invest. Ophthal. Vis. Sci. 39, 2674–2680.
16. Delany, A. M., Amling, M., Priemel, M., Howe, C., Baron, R., and Canalis, E. (2000) J. Clin. Invest. 105, 915–923.
17. Bass, A., Kugman, L. H., Samulewicz, S. J., and Howe, C. (2001) BMC Cell Biol. 2, 15.
18. Brandshaw, A. D., Reed, M. J., and Sage, H. (2002) J. Histochem. Cytochem. 50, 1–10.
19. Strandjord, T. P., Madtes, D. K., Weiss, D. J., and Sage, H. (1999) Am. J. Physiol. 277, 628–635.
20. Savani, R. C., Zhou, Z., Arguiri, E., Wang, S., Vu, D., Howe, C. C., and DeLisser, H. M. (2000) Am. J. Physiol. 279, L741–L750.
21. Johnston, I. G., Paladino, T., Gurd, J. W., and Brown, I. R. (1990) Neuron 2, 165–176.
22. Guermah, M., Crisanti, P., Laugier, D., Declee, P., Bidou, L., Pessac, B., and Calothy, G. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4503–4507.
23. Shiibahara, M., Mashimo, J., Mita, A., Kuroki, T., and Nose, K. (1993) Eur. J. Biochem. 217, 13–19.
24. Alliel, P. M., Perin, J. P., Jolles, P., and Bonnet, F. J. (1993) Eur. J. Biochem. 214, 347–350.
Characterization of SMOC-1, a Novel Modular Calcium-binding Protein in Basement Membranes

Christian Vannahme, Neil Smyth, Nicolai Miosge, Silke Gösling, Christian Frie, Mats Paulsson, Patrik Maurer and Ursula Hartmann

J. Biol. Chem. 2002, 277:37977-37986.
doi: 10.1074/jbc.M203830200 originally published online July 18, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M203830200

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

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

This article cites 65 references, 21 of which can be accessed free at http://www.jbc.org/content/277/41/37977.full.html#ref-list-1