The EMBO Journal Vol.17 No.15 pp.4249–4256, 1998

Structure, function and tissue forms of the C-terminal globular domain of collagen XVIII containing the angiogenesis inhibitor endostatin

Takako Sasaki, Naomi Fukai[1], Karlheinz Mann, Walter Göhring, Björn R.Olsen[1] and Rupert Timpl[2]

Max-Planck-Institut für Biochemie, 82152 Martinsried, Germany and [1]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

[2]Corresponding author

The C-terminal domain NC1 of mouse collagen XVIII (38 kDa) and the shorter mouse and human endostatins (22 kDa) were prepared in recombinant form from transfected mammalian cells. The NC1 domain aggregated non-covalently into a globular trimer which was partially cleaved by endogenous proteolysis into several monomers (25–32 kDa) related to endostatin. Endostatins were obtained in a highly soluble, monomeric form and showed a single N-terminal sequence which, together with other data, indicated a compact folding. Endostatins and NC1 showed a comparable binding activity for the microfibrillar fibulin-1 and fibulin-2, and for heparin. Domain NC1, however, was a distinctly stronger ligand than endostatin for sulfatides and the basement membrane proteins laminin-1 and perlecan. Immunological assays demonstrated endostatin epitopes on several tissue components (22–38 kDa) and in serum (120–300 ng/ml), the latter representing the smaller variants. The data indicated that the NC1 domain consists of an N-terminal association region (~50 residues), a central protease-sensitive hinge region (~70 residues) and a C-terminal stable endostatin domain (~180 residues). They also demonstrated that proteolytic release of endostatin can occur through several pathways, which may lead to a switch from a matrix-associated to a more soluble endocrine form.

Keywords: binding properties/collagen/proteolysis/recombinant production/serum form

Introduction

The formation of new blood vessels is crucial for tissue regeneration and for efficient tumor growth, and requires the concerted action of various angiogenic factors and inhibitors (Folkman, 1995; Hanahan and Folkman, 1996). Inhibitors have received particular attention because of their therapeutic potential, as shown in several experimental tumor models (O'Reilly *et al.*, 1996; Boehm *et al.*, 1997). A diverse group of protein inhibitors have been characterized and were identified by their ability to inhibit endothelial cell proliferation, migration and cord formation and to suppress angiogenesis and tumor growth. They include fibronectin (Homandberg *et al.*, 1985), transforming growth factor-β1 (Müller *et al.*, 1987), thrombospondin-1 (Good *et al.*, 1990; Taraboletti *et al.*, 1990a; Iruela-Arispe *et al.*, 1991), SPARC/BM-40/osteonectin (Funk and Sage, 1993), platelet factor 4 (Maione *et al.*, 1990), prolactin (Clapp *et al.*, 1993), the plasminogen-derived angiostatin (O'Reilly *et al.*, 1994) and the collagen XVIII fragment endostatin (O'Reilly *et al.*, 1997). Of particular interest was the observation that several inhibitors, such as those derived from fibronectin (Homandberg *et al.*, 1985), prolactin (Clapp *et al.*, 1993), platelet factor 4 (Gupta *et al.*, 1995) and plasminogen (O'Reilly *et al.*, 1994, 1996), require proteolytic release or activation in order to acquire inhibitory activity. However, such an activation is apparently not needed for thrombospondin-1 (Good *et al.*, 1990).

The mechanism of action of angiogenesis inhibitors on endothelial cells and their receptors is so far unknown. Studies with platelet factor 4 (Maione *et al.*, 1990) and thrombospondin (Vogel *et al.*, 1993) indicated that their heparin-binding domains were involved, suggesting competition with the angiogenic basic fibroblast growth factor (bFGF) for proteoglycan receptors. A strong affinity for heparin has also been demonstrated for angiostatin and endostatin (O'Reilly *et al.*, 1994, 1997), yet other studies showed inhibitory activity for non-heparin-binding thrombospondin fragments (Good *et al.*, 1990; Tolsma *et al.*, 1993), which may bind to the CD36 receptor (Dawson *et al.*, 1997), suggesting a complex array of biologically active sites. A similar complexity apparently exists for angiostatin, in which four individual kringle domains each have anti-proliferative activity, but to different extents (Cao *et al.*, 1996, 1997). There was apparently no correlation with lysine affinity, but whether or not there is a correlation with heparin affinity remains to be established. This diversity of angiogenesis inhibitors means that each component will require its own precise molecular analysis and it may be found that no common interaction mechanism exists.

We have recently determined the X-ray structure of mouse endostatin, demonstrating a compact globular folding and a core structure related to the carbohydrate-recognition domain of C-type lectins (Hohenester *et al.*, 1998). In the present study, we have compared its structure, binding properties and tissue forms with those of its parental collagen XVIII domain NC1. Collagen XVIII was shown by sequence analysis to consist of a central, interrupted triple-helical domain flanked at the N- (domain NC11) and C-termini (domain NC1) by larger non-triple-helical, presumably globular structures (Abe *et al.*, 1993; Oh *et al.*, 1994a; Rehn and Pihlajaniemi, 1994; Rehn *et al.*, 1994). Collagen XVIII was also shown to be expressed in a large number of tissues and was localized in vessel walls and several basement membranes using antibodies against domain NC11 (Muragaki *et al.*, 1995). Proteolytic release of NC1 and endostatin may occur

```
  1  AGQVRIWATY QTMLDKIREV PEGWLIFVAE REELYVRVRN GFRKVLLEAR
     S - L   R  A  GQVH              Q       Q       Q
       *

                 ▼▼        ▼▼
 51  TALPRGTGNE VAALQPPLVQ LHEGSPYTRR EYSYSTARPW RADDILANPP
      P     D          V     DSN  P    HPHP             S

                     ▼▼▼   ▼          ↓
101  RLPDRQPYPG VPHHHSSYVH LPPARPTLSL AHTHQDFQPV LHLVALNTPL
        EP      A   -       R     SPP   S R             S
     *          *          **                      △
                       △    △    △
                   1
151  SGGMRGIRGA DFQCFQQARA VGLSGTFRAF LSSRLQDLYS IVRRADRGSV
                              A                         AA

201  PIVNLKDEVL SPSWDSLFSG SQGQLQPGAR IFSFDGRDVL RHPAWPQKSV
             L  F   EA      E P K           K       T
                     2                                3
251  WHGSDPSGRR LMESYCETWR TETTGATGQA SSLLSGRLLE QKAASCHNSY
           N     T           APS          G    G  S     HA
        4
301  IVLCIENSFM TSFSK
                 A-
```

Fig. 1. Amino acid sequence of collagen XVIII domain NC1, which contains the endostatin structure, and identification of protease-sensitive sites. The mouse sequence (Oh *et al.*, 1994a) is shown in the top line and human substitutions (Oh *et al.*, 1994b) underneath. A dash indicates a deletion. Cysteines are consecutively numbered 1–4 on top. The N-terminal cleavage site of mouse tumor endostatin (O'Reilly *et al.*, 1997) is marked by an arrow. Arrowheads mark endogenous cleavage sites in mouse NC1 (▼) and in human plasma forms (△) including position 141/142 reported by Ständker *et al.* (1997). Thick lines underline sequences of mouse endostatin where the entire structure was determined by Edman degradation and mass spectrometry demonstrating the disulfide connections Cys1–Cys4 and Cys2–Cys3 (see text). Thin lines show further sequences determined by Edman degradation of single-chain peptides. Five sequence corrections of human NC1 are marked by asterisks.

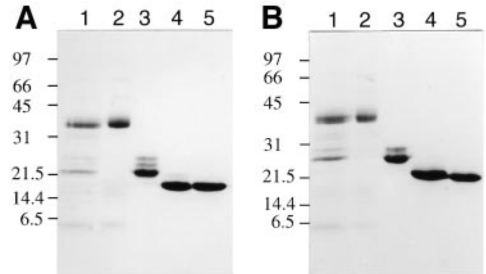

Fig. 2. SDS gel electrophoresis of purified collagen XVIII NC1 domain and endostatins under non-reducing (**A**) and reducing (**B**) conditions. Lanes were loaded with mouse NC1 after heparin chromatography (lane 1) and after separation (see Figure 3) into the NC1-38 (lane 2) and NC1-25 (lane 3) components and with mouse (lane 4) and human (lane 5) endostatin. The slight shift to a slower mobility after reduction indicates opening of internal disulfide bonds. Positions of globular markers (in kDa) are denoted in the left margins.

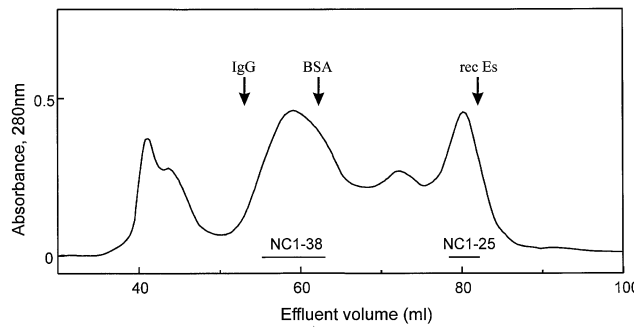

Fig. 3. Molecular sieve chromatography separation of oligomeric NC1 (NC1-38) from a monomeric fragment NC1-25. The two pools collected (horizontal bars) are shown by their electrophoresis patterns in Figure 2 (lanes 2 and 3). The Superose 12 column was run in 0.05 M Tris–HCl pH 7.4, 0.5 M NaCl and calibrated with immunoglobulin G (IgG), bovine serum albumin (BSA) and recombinant mouse endostatin (rec Es).

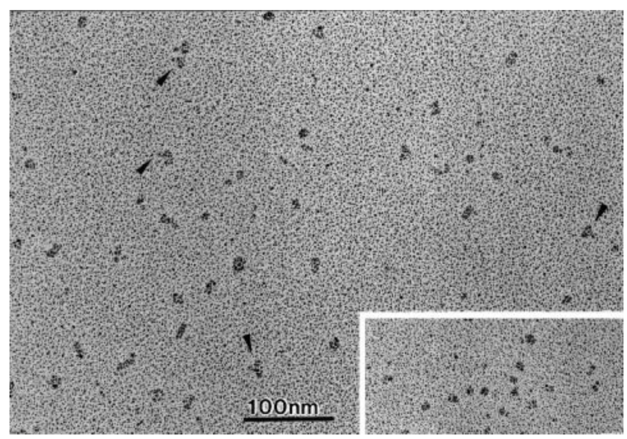

Fig. 4. Electron microscopy of oligomeric domain NC1 (NC1-38) after rotary shadowing. Note the abundance of globular and some elongated structures which were distinctly larger than the globular domain G3 of nidogen with the same monomeric molecular mass of 38 kDa (see insert in the lower right corner). A few NC1-38 particles are extended by a short stalk (marked by arrowheads) which may be obscured in other cases due to adsorption to the mica disc.

through variable mechanisms both in cell cultures and tissues. This can now be explained from the structure of trimeric NC1, which consists of a region responsible for oligomerization, a protease-labile segment and the active endostatin domain. This processing also changes the binding activity for various extracellular ligands.

Results

Production and structural characterization of endostatin and domain NC1

The C-terminal domain NC1 of collagen XVIII consists of a non-triple-helical sequence of 315 (mouse) or 312 (human) residues (Figure 1). Endostatin previously obtained from murine hemangioendothelioma medium was shown to start at position 132 and, because of its size (20 kDa), must cover the complete C-terminal part of NC1 (O'Reilly *et al.*, 1997). We have now made episomal expression vectors for mouse (positions 130–315) and human (129–312) endostatins as well as mouse NC1 (1–315) and used them to transfect human kidney 293-EBNA cells. A radioimmunoassay analysis (see below) of serum-free medium from cultured cells demonstrated a high production of both endostatins (6–20 µg/ml/day) and a somewhat lower production of NC1. Both endostatins bound efficiently to a heparin–Sepharose column and were, after elution with NaCl and molecular sieve chromatography, already of high purity, as shown by electrophoresis (Figure 2). The apparent molecular mass was ~22 kDa. Recombinant NC1 also bound to heparin and was eluted as two major bands of 25 and 38 kDa, together with some minor bands (Figure 2). The two major bands could be separated on a Superose 12 column, which demonstrated that the intact form has a molecular mass of ~100 kDa (Figure 3). This indicated that the 38 kDa band associates non-covalently to form a trimer, while the 25 kDa band was monomeric like endostatin. Oligomerization was also demonstrated by electron microscopy, which revealed globular particles that were mostly larger than a typical globular domain of 38 kDa (Figure 4).

Interestingly, some of the NC1 structures appeared asymmetric, having an additional short stalk.

Mouse endostatin, as expected, showed a single N-terminal sequence APLAHTHQ, the AP being artificially introduced by its insertion next to a signal peptide sequence in the expression vector. The human endostatin sequence was APLAHSHRDF, the APL being foreign. The 38 kDa NC1 N-terminus was APLAGQVR, while the majority of the 25 kDa fragment started at position 117, with two minor variants that were one or two residues shorter (Figure 1). Hexosamine analysis of endostatin and NC1 failed to show significant amounts (<0.5 residues per molecule). The flow-through fraction after heparin chromatography was used to search for variants modified by glycosaminoglycans, using DEAE–cellulose chromatography and monitoring by a carbazol assay (Costell *et al.*, 1997). No such variants were found, despite the presence of several SG sequences in endostatin (Figure 1), which could be potential glycosaminoglycan acceptor sites (Oh *et al.*, 1994a).

Conformation and protease stability

Far-UV circular dichroism (CD) spectroscopy revealed a similar spectrum for both endostatins, with a characteristic minimum (θ = –15 000 deg.cm^2.dmol^{-1}) at 210 nm (data not shown). Secondary structure estimates indicated ~10% α-helix and ~70% β-structure. Digestion of mouse endostatin with plasmin, trypsin, α-chymotrypsin, pancreatic elastase or endoproteinase Glu-C followed by electrophoresis prior to or after reduction revealed no or only minor changes in the band patterns (not shown). Three digests were subjected directly to Edman degradation, which revealed some minor sequences in addition to the major N-terminal sequence. This indicated a very limited cleavage at some peptide bonds between positions 240 and 291 (data not shown) and thus a compact folding of endostatin at neutral pH, as also demonstrated by X-ray crystallography (Hohenester *et al.*, 1998).

Exposure of mouse endostatin to pepsin at pH 3, however, caused complete digestion within 4 h to small peptides which no longer bound to heparin. Reverse-phase chromatography resolved this digest into at least 18 major peaks, all of which were analyzed by Edman degradation. Two of these peaks contained a double sequence, encompassing either the region containing Cys1 and Cys4 or that containing Cys2 and Cys3. Mass spectrometry revealed mainly dimers of 3497.57 Da (Cys2,3 peak) or 1939.26 Da (Cys1,4 peak), demonstrating that these cysteines are disulfide-linked (Figure 1).

The NC1 domain was distinctly more sensitive to endogenous proteolysis (see above) and omission of protease inhibitors during purification caused the appearance of electrophoretic bands of 30–32 kDa and 5 kDa in significant amounts. The 30 kDa component started at four different sites (positions 63–72; Figure 1) while the 5 kDa fragment showed the original N-terminus. The latter fragment did not bind to heparin any more and was shown to elute at a position of ~12 kDa in molecular sieve chromatography, indicating oligomerization.

Binding to carbohydrate and protein ligands

A small heparin HiTrap column was used for a closer analysis of binding parameters. Both of the endostatins and NC1 eluted at the same NaCl concentration (0.34 M). Mouse endostatin contains 15 Arg and four Lys residues (Figure 1) which could contribute to binding. After specifically modifying the Lys side chains, the same efficient heparin binding (>99%) and elution at 0.33 M NaCl was observed. After modification of Arg, however, a major fraction (83%) eluted in overlapping peaks at 0.08–0.2 M NaCl, while 3% did not bind and 14% eluted at 0.34 M NaCl. This suggested that several Arg residues are involved in binding and the heterogeneous binding population was probably due to their being modified to different extents. In addition, trimeric mouse NC1 bound to sulfatides almost as efficiently as laminin-1. Mouse and human endostatin were, however, less active by a factor of 100–200 (Figure 5). A 10-fold weight excess of soluble heparin caused no or only slight inhibition of sulfatide binding, indicating that the two interactions depend on different binding epitopes.

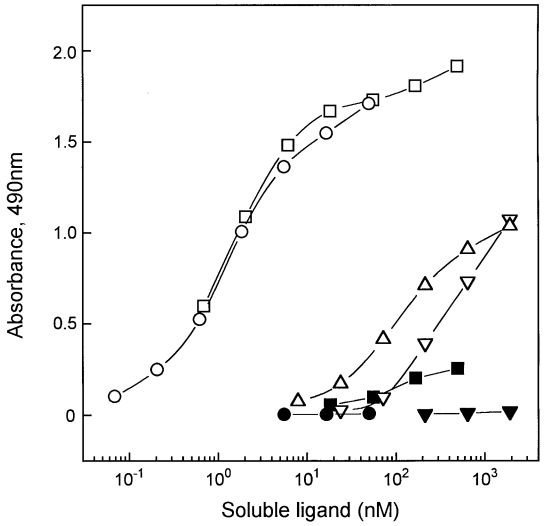

Fig. 5. Binding of NC1 and endostatins to immobilized sulfatides in solid-phase assay. Soluble ligands were laminin-1 complexed to nidogen (○), mouse (▽) and human (△) endostatin and trimeric mouse NC1 (□). Corresponding solid symbols show background binding in the absence of sulfatides.

The binding of several extracellular matrix proteins to immobilized trimeric NC1 and mouse endostatin was determined in solid-phase assays. NC1 showed a strong interaction with perlecan and laminin-1 (half-maximal binding at 0.1–1 nM) while endostatin was again 100-fold less active (Figure 6). A strong and comparable binding of both NC1 and endostatin was observed with fibulin-1 and -2. This was confirmed in a kinetic binding assay (Table I) which indicated K_d values for the interaction with fibulin-1 in the range 20–100 nM and for fibulin-2 in the range 17–34 nM. This binding occurred with similar affinities also in the presence of EDTA, unlike other fibulin interactions, which are strictly dependent on calcium (Sasaki *et al.*, 1995). Other extracellular proteins such as nidogen-1 and BM-40 were only very weak or inactive ligands for NC1 and endostatin, underscoring the specificity of the other interactions.

Nature of endostatin in tissues and serum

Rabbit antisera were raised against mouse and human endostatin and used for immunochemical assays of various

biological samples and for immunohistochemistry. The antisera showed similar high titers for both endostatins in ELISA and radioimmunoassays. This allowed us to establish sensitive radioimmunoinhibition assays, as shown for mouse endostatin in Figure 7. Mouse endostatin and NC1 inhibited in a very similar fashion, with as little as 5 ng/ml required for a half-maximal effect. Reduction and alkylation of mouse endostatin reduced its inhibitory activity to low levels, indicating that the antiserum recognizes mainly conformation-dependent epitopes. Furthermore, inhibition profiles of mouse serum and tissue extracts showed identical gradients to that of endostatin (Figure 7), indicating that they share the same set of epitopes. A similar assay specific for human endostatin showed a comparable sensitivity for detecting the recombinant protein and similar material in biological samples (data not shown).

A quantitative analysis by these assays demonstrated an endostatin content of 120 ± 12 ng/ml in the serum of adult mice and 22 ± 13 ng/ml in their urine. Concentrations of 120–300 ng/ml were determined for six sera and one plasma obtained from healthy human donors. Endostatin could not be detected (<2 ng/ml) in conditioned medium from 10 different human and mouse cell lines including fibroblasts, epithelial and tumor cells. Low levels (5–15 ng/ml) were detected in the medium of human aortic and microvascular endothelial cells and that of F9 teratocarcinoma cells. Tissue extracts (1 ml per 0.1–0.2 g wet tissue) from mouse brain, skeletal muscle, heart, kidney, testis and liver showed increasing concentrations of endostatin in the range 0.3–2 μg/ml. These amounts exceeded that of serum endostatin by at least a factor of 10, indicating that endostatin is generated in the tissues and not derived from blood contamination. In immunoblots of these extracts, bands with a mobility similar to NC1 were observed to a variable degree, but there was none with the mobility of recombinant endostatin (Figure 8A). Bands with an intermediate electrophoretic mobility were more prominent in some of the extracts, however. Testis and skeletal muscle contained additional bands of ~100 kDa, which are presumably too short to represent full-length α1(XVIII) collagen chains (Muragaki *et al.*, 1995).

More than 90% of the endostatin antigen present in mouse and human serum and in mouse liver extract bound to a heparin HiTrap column and most of it was eluted as a narrow peak at 0.34 M NaCl. A semi-preparative isolation from human plasma (see Materials and methods) showed three reactive bands of 23–26 kDa by immunoblotting (Figure 8B). These bands could be sufficiently separated from large amounts of contaminating proteins to provide a clear identification by Edman degradation. This demonstrated three cleavage sites between positions 116–126 of the human NC1 sequence (Figure 1).

Immunohistochemistry with affinity-purified polyclonal antibodies demonstrated specific staining mainly in the

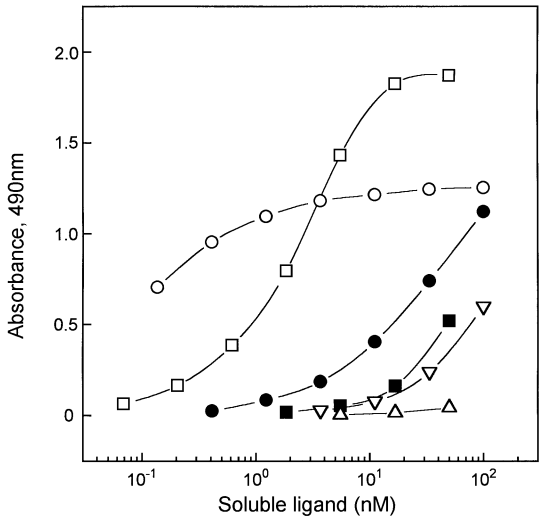

Fig. 6. Solid-phase assay of the binding of mouse endostatin (solid symbols) and trimeric domain NC1 (open symbols) to various basement membrane proteins. Endostatin and NC1 were used as immobilized ligands. Soluble ligands were laminin-1 complexed to nidogen (□, ■) and perlecan (○, ●) and their binding was detected by specific antibodies. In addition, background binding to the serum albumin used for blocking is shown for laminin-1 (△) and perlecan (▽).

Table I. Equilibrium dissociation constants of the binding of fibulins to endostatin and NC1

Fibulin ligands[a]	K_d values[b] (nM) of binding to	
	Endostatin	NC1-35
Fibulin-1C (S)	100 ± 30 (I)	111 ± 12 (I)
Fibulin-1C (I)	46 ± 7 (S)	24 ± 9 (S)
Fibulin-2 (S)	33 ± 12 (I)	17 ± 6 (I)
Fibulin-2 (I)	ND[c] (S)	34 ± 16 (S)

[a]Each ligand was used in either soluble (S) or immobilized (I) form in surface plasmon resonance assays.
[b]Values represent the mean and error range of 2–3 determinations.
[c]Questionable evaluation because of low uptake of resonance units.
ND, not detected.

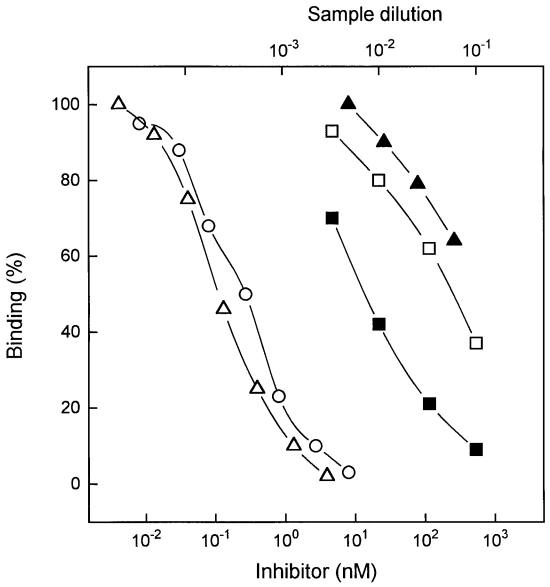

Fig. 7. Radioimmunoinhibition assay profiles for the analysis of mouse endostatin. The assay consisted of fixed concentrations of ^{125}I-labeled mouse endostatin (1 ng) and a corresponding antiserum (diluted 1:2000). Purified inhibitors examined were mouse endostatin (○), trimeric mouse domain NC1 (△) and reduced and alkylated mouse endostatin (▲) at the concentrations indicated. Mouse serum (□) and heart extract (■) were used at the sample dilutions shown at the top of the figure.

basement membranes zones of mouse embryos and adult mouse tissues. In developing embryos strong staining was observed in skin, brain and vascular basement membranes (Figure 9A). Basement membranes of developing kidney, lung and choroid plexus also showed positive staining. In adult tissues, basement membranes, liver cells (especially around central veins) and kidney cortex showed strong staining (data not shown). Interestingly, in the placenta the fetal, but not the maternal, portion showed staining with the antibody (Figure 9B). Staining of adult aorta included the intima and media, suggesting also a localization of endostatin outside of basement membrane zones.

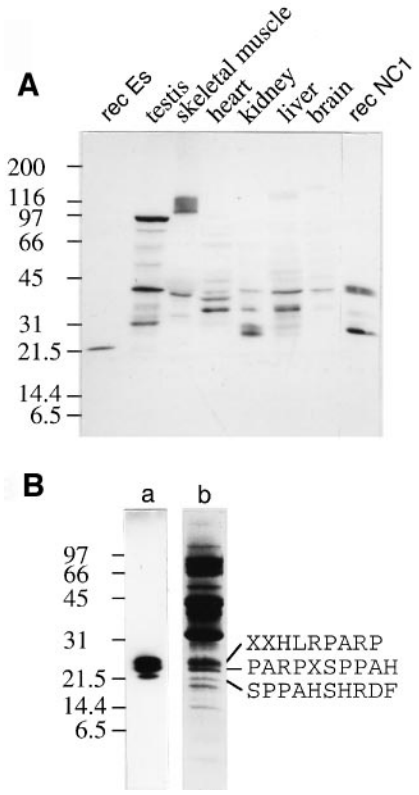

Fig. 8. Immunoblot detection of endostatin in mouse tissue extracts (**A**) and electrophoretic and sequence analysis of prominent endostatin components in human plasma (**B**). The amounts in each lane in (A) were adjusted to 1 ng endostatin as determined by radioimmunoassay. Calibration was with recombinant mouse endostatin (rec ES) and NC1 (rec NC1) and with protein markers, whose kDa values are shown. In (B), the immunoblot of a serum pool enriched for endostatin (lane a) and the corresponding Coomassie blue stain (lane b) are shown. Edman degradation of bands in the latter blot identified three endostatin bands by their N-terminal sequences (for positions, see Figure 1). X denotes an unidentified residue.

Discussion

The study established the efficient production of mouse and human endostatin using a modified episomal expression vector in 293 human kidney cells (Kohfeldt *et al.*, 1997). The production rates and yields after a two-step purification were almost as high as for recombinant endostatin obtained from bacterial transfectants and distinctly better than that of endostatin obtained from insect cells (O'Reilly *et al.*, 1997). Our recombinant endostatin was highly soluble in neutral buffer, in contrast to the bacterial product. This is consistent with a proper folding of endostatin (Hohenester *et al.*, 1998), as demonstrated here by CD spectroscopy, proteolytic resistance, nature of immunological epitopes and the efficient formation of two disulfide bridges between Cys1–Cys4 and Cys2–Cys3. The latter is in agreement with the crystal structure of mouse endostatin (Hohenester *et al.*, 1998), but at variance with a Cys1–Cys3 connection determined by mass spectrometry for human plasma-derived endostatin (Ständker *et al.*, 1997). Together, these data demonstrate that endostatin represents an autonomously folding unit within the NC1 domain of collagen XVIII. A related structure (54% sequence identity) exists at the same location in collagen XV (Oh *et al.*, 1994a; Rehn and Pihlajaniemi, 1994) and may have the same autonomous folding property. The mouse endostatin used here was also a strong inhibitor of a subcutaneous gliosarcoma in rats by reducing angiogenesis (D.Soerensen, R.Bjerkvig, R.Timpl, T.Sasaki, P.O.Iversen, H.B.Benestad and B.R.Olsen, unpublished results) and was thus comparable in its biological properties to endostatin derived from bacteria (Boehm *et al.*, 1997; O'Reilly *et al.*, 1997).

A major aim of the study was to examine the structure of collagen XVIII domain NC1 and its proteolytic cleavage to endostatin. Recombinant fragment NC1 was partially sensitive to endogenous proteolysis, but a substantial

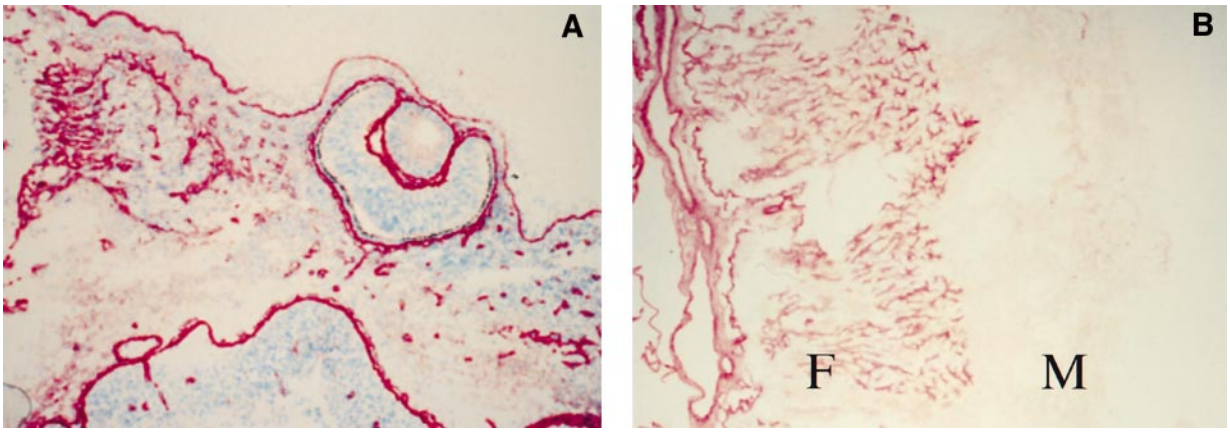

Fig. 9. Immunohistochemical localization of endostatin epitopes in normal mouse tissues. (**A**) Strong staining in skin, developing brain and vascular basement membranes of day 12 mouse embryo. (**B**) Placenta of a day 17 pregnant mouse. Note the prominent staining of vessels in the fetal portion of the placenta (F) but no staining in the maternal portion (M).

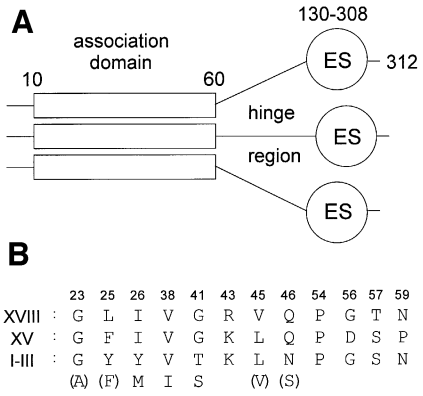

Fig. 10. Schematic model of the domain structure of collagen XVIII NC1 (**A**) and conservation of amino acids in the association domain (**B**). The domain structure includes an association domain responsible for non-covalent trimerization (residues ~10–60) and the compact endostatin domain (ES, residues 130–308) connected by a hinge region, as defined by sites sensitive to proteolytic cleavage (see Figure 1). (B) identifies positions (numbering of human α1XVIII NC1, Figure 1) conserved in collagen XVIII, the related collagen XV and the fibrillar collagens I–III. The bottom line indicates minor (in parentheses) or major deviations from the consensus sequences of collagens I–III shown above. All sequences are of human origin.

portion could be recovered as an intact 38 kDa component which assembled non-covalently into a trimeric structure. To our knowledge, this is the first direct evidence that collagen α1(XVIII) chains can form homotrimers. Based on structural data, we propose that NC1 consists of three different segments (Figure 10A) which include the globular 180 residue endostatin domains, each of 3 nm diameter (Hohenester *et al.*, 1998), at the C-terminus and the association domain of presumably not more than 50 residues at the N-terminus. These two domains are connected by a hinge region which is characterized by several sites sensitive to endogenous proteolysis. This domain model fits the gene structure remarkably well (Rehn *et al.*, 1996), with exons 38 and 39 encoding the association domain, exon 40 the hinge region, and three more exons the endostatin domain. The association domain is apparently stable, as shown by the isolation of an oligomerized 5 kDa fragment, and could correspond to the short stalks seen for several NC1 particles by electron microscopy. Interestingly, this short sequence contains 10–12 residues which are reasonably well conserved in the related collagens XV and XVIII but also in the fibrillar collagens I–III (Figure 10B). It is tempting to speculate that some of them may be crucial for self assembly, though this region is different from another type-specific self-assembly site indicated for the C-propeptide of the collagen α1(III) chains (Lees *et al.*, 1997).

A surprising observation was the broad occurrence of endostatin-like proteins in tissues and serum with a complexity in size ranging from endostatin (22 kDa) to NC1 monomer (38 kDa). Sequence analysis of endogenous recombinant fragments and of endostatins from human plasma have now identified 11 different positions in the hinge region of mouse and human NC1 that are sensitive to proteolysis (Figure 1). This indicates that several proteolytic pathways may exist for the generation of endostatin in tissues, a finding which now needs further examination. The concentration (100–300 ng/ml) of endostatin in serum (plasma) was similar to the concentrations which efficiently inhibited endothelial cell proliferation *in vitro* (O'Reilly *et al.*, 1997). This suggests that circulating forms may be involved in the homeostatic control of angiogenesis. Three serum variants slightly larger than recombinant human endostatin could be detected by sequence analysis and we assume that they all have inhibitory potential. We detected none of the smaller inactive variant recently isolated from human plasma hemofiltrate (Ständker *et al.*, 1997) and only small amounts of NC1-like proteins in human plasma. Because of its high content, liver could be one source of the circulating endostatins, but this would require a substantial proteolytic conversion of its major variants which are similar in size to NC1 (Figure 8A).

Of particular interest was the binding potential of endostatin and/or NC1 for several ubiquitous extracellular matrix proteins as indicated by *in vitro* assays. They share similar affinities for fibulin-1 and -2 which are known to associate with fibronectin and, in the case of fibulin-2, with fibrillin into tissue microfibrils (Reinhardt *et al.*, 1996; Sasaki *et al.*, 1996) that are not necessarily localized to basement membrane zones. Yet, both fibulins are also typical components of vascular structures (Pan *et al.*, 1993) as also shown previously by immunohistochemistry for the NC11 domain of collagen XVIII (Muragaki *et al.*, 1995) and now here for endostatin (Figure 9). Evidence for the specificity of the latter antibodies for epitopes in collagen XVIII and lack of cross-reaction with the most closely related collagen XV comes from the observation that no staining was observed with tissues obtained from homozygous collagen XVIII knock-out mouse embryos (N.Fukai and B.R.Olsen, unpublished results). Binding activities unique to the NC1 domain included several basement membrane ligands such as laminin-1 and perlecan. This could reflect the multivalency of trimeric NC1 or may be a specific property of its association domain or the hinge region.

Immunoblotting demonstrated that at least part of nearly full-length NC1 is released from the parental collagen in many tissues but could still be retained in the supramolecular assembly of basement membranes through non-covalent interactions with other ligands. A second proteolytic step may be required to release soluble endostatin, which could then diffuse away and act in a paracrine or endocrine fashion. A strong association with fibulins, however, may interfere with endostatin release and we speculate that subanatomical localization could be an important factor that controls availability of tissue-bound endostatin activity for angiogenesis inhibition.

The NC1 domain and endostatin shared in addition a strong binding to heparin. Since the small association domain of NC1 did not bind to heparin, this property and possibly cell inhibition may be exclusively located in the endostatin domain. The structure of mouse endostatin (Hohenester *et al.*, 1998) demonstrated that 11 out of the total 15 arginines are clustered together and some of them, as shown here, may participate in the heparin-binding epitope. As discussed above, this binding site may compete with basic FGF for cellular proteoglycan receptors and thus cause inhibition of proliferation, but alternative mechanisms cannot be excluded. Other possibilities include association with sulfatides or interaction with

specific endothelial cell receptors which have not yet been identified.

In summary, our data indicate a complex functional repertoire for domain NC1 of collagen XVIII. Besides regulation of endothelial cell proliferation as shown for endostatin (O'Reilly *et al.*, 1997), this includes oligomerization, which could be crucial for triple-helix formation, and distinct binding activities for heparin, sulfatides and several extracellular matrix proteins. The precise mapping of these activities by site-directed mutagenesis now seems feasible based on recombinant production and the X-ray structure of endostatin.

Materials and methods

Sources of proteins and other components

Various basement membrane proteins either purified from a mouse tumor or obtained in recombinant form were those used in previous studies (Sasaki *et al.*, 1995). Sulfatides were purchased from Sigma.

Construction of expression vectors

Mouse α1(XVIII) cDNA clone mc3b (Oh *et al.*, 1994a) and human α1(XVIII) cDNA clone pNF18-2 (Oh *et al.*, 1994b) were used as templates to amplify the sequences encoding NC1 and endostatin by polymerase chain reaction (PCR) with Vent polymerase (New England Biolabs) following the manufacturer's instructions and using the following oligonucleotide primers:

M-1: GTCAGCTAGCTCATACTCATCAGGAC;
M-2: GTCAGCTAGCTGGGCAGGTGAGGATC;
M-3: GTCACTCGAGCTATTTGGAGAAAGAGGTC;
H-1: GTCAGCTAGCCCACAGCCACCGCG;
H-2: GTCACTCGAGCTACTTGGAGGCAGTCATG.

The primers M-1 and M-3 were used for mouse endostatin, M-2 and M-3 for mouse NC1, and H-1 and H-2 for human endostatin. In addition to the annealing sequences, these primers contained an *Nhe*I site at the 5′ end (M-1, M-2, H-1) or a stop codon followed by a *Xho*I site at the 3′ end (M-3, H-2) in order to allow the constructs to be inserted in-frame with the BM-40 signal peptide (Mayer *et al.*, 1993). The sequences of the PCR fragments were confirmed by cycle sequencing using Dye Terminator Cycle Sequencing Ready Reaction Kit (ABI). All PCR products were restricted with *Nhe*I and *Xho*I and cloned into the corresponding sites of the modified episomal expression vector pCEP-Pu (Kohfeldt *et al.*, 1997).

Expression and purification of recombinant and plasma-derived proteins

Human embryonic kidney cells which express the EBNA-1 protein from Epstein–Barr virus (293-EBNA cells; Invitrogen) were transfected with the expression vectors (Kohfeldt *et al.*, 1997). Transfected cells were selected with puromycin (0.5 μg/ml) and used to collect serum-free conditioned medium. EDTA (2 mM), *N*-ethylmaleimide (0.5 mM) and phenylmethylsulfonyl fluoride (0.5 mM) were added to the NC1 medium in order to reduce proteolysis. The medium (~1 l) was dialyzed against 0.1 M NaCl, 0.05 M Tris–HCl, pH 7.4 and then applied to a heparin–Sepharose Cl-6B column (Pharmacia, 2.5×30 cm) which was equilibrated in the same buffer. Elution was with a linear 0.1–1.0 M NaCl gradient (500 ml). Both endostatins eluted at 0.3–0.4 M NaCl and mouse NC1 eluted at 0.35–0.45 M NaCl. The endostatins were concentrated and further purified on a Superose 12 column (HR16/50, Pharmacia) equilibrated in 0.2 M ammonium acetate, pH 6.8. NC1 was purified on the same Superose column equilibrated in 0.05 M Tris–HCl, pH 7.4, 0.5 M NaCl in order to prevent NC1 precipitation at low ionic strength. The purified components were kept in the buffer used for chromatography and stored at –20°C.

Purification from human plasma (300 ml) was initially carried out on heparin–Sepharose followed by a 5 ml heparin HiTrap (Pharmacia) column and elution was monitored by radioimmunoassay. The enriched plasma proteins were then subjected to reverse-phase HPLC on a C4 column, from which endostatin antigens eluted mainly in a single peak at 23% acetonitrile.

Methods of protein characterization

Hydrolysis (16 h, 110°C) with either 6 M or 3 M HCl was used to determine the amino acid and hexosamine contents, respectively, on a Biotronik LC 3000 analyzer. SDS–polyacrylamide gel electrophoresis in gradient gels (10–20%) followed established protocols. Edman degradation on Procise and 473A sequencers (Applied Biosystems) was performed following the manufacturer's instructions. Electrophoretically separated protein bands were blotted onto Immobilon PSQ membranes (Millipore) for sequencing. Small peptides were separated by reverse-phase chromatography on a C18 column. Digestions with neutral proteinases in 0.05 M Tris–HCl, pH 7.4, 2 mM $CaCl_2$ (24 h, 37°C) and with pepsin in 0.1 M acetic acid (4 h, 37°C) were performed at enzyme:substrate ratios of 1:100. Side chain modification of Lys by selective acetylation (Fraenkel-Conrat, 1957) and of Arg by phenylglyoxal (Takahashi, 1968) followed established procedures. CD spectra were measured and evaluated as described previously (Schulze *et al.*, 1995). Matrix-assisted laser desorption ionization (MALDI) mass spectrometry was performed following a previously reported method (Mann *et al.*, 1996).

Binding assays

Solid-phase binding assays with plastic immobilized endostatin and NC1 and with soluble ligands followed a previous procedure (Sasaki *et al.*, 1995). A similar binding assay with coated sulfatides (10 μg/well) has been described (Taraboletti *et al.*, 1990b). A 1 ml heparin HiTrap column equilibrated in 0.05 M Tris–HCl, pH 7.4 was used for analytical affinity chromatography of protein samples (0.2–0.3 mg) and was eluted with a linear gradient of 0–1 M NaCl. The content of different pools was determined by amino acid analysis or radioimmunoassay.

Surface plasmon resonance assays were performed with BIAcore instrumentation (BIAcore AB, Uppsala) using various ligands immobilized to CM-5 sensor chips. Binding assays were carried out at 25°C in neutral buffer containing either 1 mM $CaCl_2$ or 4 mM EDTA (Sasaki *et al.*, 1998). Dissociation rate constants were calculated from association and dissociation curves according to the 1:1 model following the manufacturer's instructions (BIAevaluation software version 2.1). Various other experimental details to assure that an 1:1 dissociation model is applicable were those described previously (Sasaki *et al.*, 1998).

Immunological assays

Immunization of rabbits, affinity purification of antibodies and radioimmunoassays followed established protocols (Timpl, 1982). Tissues of 4-week-old Balb/c mice were extracted (1 ml/100–200 mg wet tissue) with neutral buffer containing protease inhibitors and 1% NP-40, 0.1% SDS and 0.5% sodium deoxycholate and used for immunoblotting with affinity-purified antibodies (Sasaki *et al.*, 1996) and for radioimmunoassays. Serum and urine was collected from female 6- to 8-week-old Balb/c mice ($n = 13$) for radioimmunoassays.

Immunohistochemistry with the affinity-purified anti-mouse endostatin polyclonal antibody was performed with embryonic and adult mouse tissues, including placenta. Tissues were fixed in 4% paraformaldehyde in 0.1 M phosphate buffer (PB) pH 7.4 at 4°C overnight. After immersion in 20% sucrose in PB to prevent freezing artifacts, the tissues were embedded in OCT compound (Miles, USA). Frozen tissue sections were cut at a thickness of 8 μm. The sections were treated with 0.2 M HCl for 10 min, washed twice in phosphate-buffered saline (PBS) and incubated with Protease Type XXIV (Sigma F-4658; 125 μg/ml in H_2O) at 37°C for 10 min. Tissues were washed twice with PBS for 5 min each, and primary antibody was applied for 2 h at room temperature. The specimens were washed twice with PBS for 5 min, and secondary antibody (Sigma; anti rabbit IgG alkaline phosphatase-conjugated, 1:300 dilution in PBS) was applied for 30 min at room temperature. After washing the samples three times with PBS (5 min each), the sections were stained by reaction with fast-red (Sigma, Fast-red TR/Naphthol AS-MX). The slides were counterstained with 0.2% methyl green, mounted and observed. Controls were sections without the primary antibody staining and sections stained with rabbit preimmune sera, which were all negative.

Acknowledgements

The authors acknowledge the expert technical assistance of Vera van Delden, Mischa Reiter, Christa Wendt, Hanna Wiedemann and Albert Ries and to Dr C.Eckerskorn for mass spectrometry. The study was supported by EC grant BIO4-CT96-0537 and NIH grant AR 36820.

References

Abe,N., Muragaki,Y., Yoshioka,H., Inoue,H. and Ninomiya,Y. (1993) Identification of a novel collagen chain represented by extensive

interruptions in the triple-helical region. *Biochem. Biophys. Res. Commun*., **196**, 576–582.
Boehm,T., Folkman,J., Browder,T. and O'Reilly,M.S. (1997) Anti-angiogenic therapy of cancer does not induce acquired drug resistance. *Nature*, **390**, 404–407.
Cao,Y. *et al*. (1996) Kringle domains of human angiostatin: characterization of the anti-proliferative activity on endothelial cells. *J. Biol. Chem*., **271**, 29461–29467.
Cao,Y., Chen,A., An,S.S.A., Ji,R.-W., Davidson,D., Cao,Y. and Llinas,M. (1997) Kringle 5 of plasminogen is a novel inhibitor of endothelial cell growth. *J. Biol. Chem*., **272**, 22924–22928.
Clapp,C., Martial,J.A., Guzman,R.C., Rentier-Delrue,F. and Weiner,R.I. (1993) The 16-kilodalton N-terminal fragment of human prolactin is a potent inhibitor of angiogenesis. *Endocrinology*, **133**, 1292–1299.
Costell,M., Mann,K., Yamada,Y. and Timpl,R. (1997) Characterization of recombinant perlecan domain I and its substitution by glycosaminoglycans and oligosaccharides. *Eur. J. Biochem*., **243**, 115–121.
Dawson,D.W., Pearce,S.F.A., Zhong,R., Silverstein,R.L., Frazier,W.A. and Bouck,N.P. (1997) CD36 mediates the *in vitro* inhibitory effects of thrombospondin-1 on endothelial cells. *J. Cell Biol*., **138**, 707–717.
Folkman,J. (1995) Clinical applications of research on angiogenesis. *N. Engl. J. Med*., **333**, 1757–1763.
Fraenkel-Conrat,H. (1957) Methods for investigating the essential groups for enzyme activity. *Methods Enzymol*., **4**, 247–269.
Funk,S.E. and Sage,E.H. (1993) Differential effects of SPARC and cationic SPARC peptides on DNA synthesis by endothelial cells and fibroblasts. *J. Cell. Physiol*., **154**, 53–63.
Good,D.J., Polverini,P.J., Rastinejad,F., Le Beau,M.M., Lemons,R.S., Frazier,W.A. and Bouck,N.P. (1990) A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin. *Proc. Natl Acad. Sci. USA*, **87**, 6624–6628.
Gupta,S.K., Hassel,T. and Singh,J.P. (1995) A potent inhibitor of endothelial cell proliferation is generated by proteolytic cleavage of the chemokine platelet factor 4. *Proc. Natl Acad. Sci. USA*, **92**, 7799–7803.
Hanahan,D. and Folkman,J. (1996) Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell*, **86**, 353–364.
Hohenester,E., Sasaki,T., Olsen,B.R. and Timpl,R. (1998) Crystal structure of the angiogenesis inhibitor endostatin at 1.5 Å resolution. *EMBO J*., **17**, 1656–1664.
Homandberg,G.A., Williams,J.E., Grant,D., Schumacher,B. and Eisenstein,R. (1985) Heparin-binding fragments of fibronectin are potent inhibitors of endothelial cell growth. *Am. J. Pathol*., **120**, 327–332.
Iruela-Arispe,M.L., Bornstein,P. and Sage,H. (1991) Thrombospondin exerts an antiangiogenic effect on cord formation by endothelial cells *in vitro*. *Proc. Natl Acad. Sci. USA*, **88**, 5026–5030.
Kohfeldt,E., Maurer,P., Vannahme,C. and Timpl,R. (1997) Properties of the extracellular calcium binding module of the proteoglycan testican. *FEBS Lett*., **414**, 557–561.
Lees,J.F., Tasab,M. and Bulleid,N.J. (1997) Identification of the molecular recognition sequence which determines the type-specific assembly of procollagen. *EMBO J*., **16**, 908–916.
Maione,T.E., Gray,G.S., Petro,J., Hunt,A.J., Donner,A.L., Bauer,S.I., Carson,H.F. and Sharpe,R.J. (1990) Inhibition of angiogenesis by recombinant human platelet factor-4 and related peptides. *Science*, **247**, 77–79.
Mann,K., Mechling,D.E., Bächinger,H.P., Eckerskorn,C., Gaill,F. and Timpl,R. (1996) Glycosylated threonine but not 4-hydroxyproline dominates the triple helix stabilizing positions in the sequence of a hydrothermal vent worm cuticle collagen. *J. Mol. Biol*., **261**, 255–266.
Mayer,U., Nischt,R., Pöschl,E., Mann,K., Fukuda,K., Gerl,M., Yamada,Y. and Timpl,R. (1993) A single EGF-like motif of laminin is responsible for high affinity nidogen binding. *EMBO J*., **12**, 1879–1885.
Müller,G., Behrens,J., Nussbaumer,U., Böhlen,P. and Birchmeier,W. (1987) Inhibitory action of transforming growth factor beta on endothelial cells. *Proc. Natl Acad. Sci. USA*, **84**, 5600–5604.
Muragaki,Y., Timmons,S., Griffith,C.M., Oh,S.P., Fadel,B., Quertermous,T. and Olsen,B.R. (1995) Mouse Col18a1 is expressed in a tissue-specific manner as three alternative variants and is localized in basement membrane zones. *Proc. Natl Acad. Sci. USA*, **92**, 8763–8767.
Oh,S.P., Kamagata,Y., Muragaki,Y., Timmons,S., Ooshima,A. and Olsen,B.R. (1994a) Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-XaaYaa repeats identify a distinct family of collagenous proteins. *Proc. Natl Acad. Sci. USA*, **91**, 4229–4233.
Oh,S.P., Warman,M.L., Seldin,M.F., Cheng,S.-D., Knoll,J.H.M., Timmons,S. and Olsen,B.R. (1994b) Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localization of the α1(XVIII) collagen gene to mouse chromosome 10 and human chromosome 21. *Genomics*, **19**, 494–499.
O'Reilly,M.S. *et al*. (1994) Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell*, **79**, 315–328.
O'Reilly,M.S., Holmgren,L., Chen,C.C. and Folkman,J. (1996) Angiostatin induces and sustains dormancy of human primary tumors in mice. *Nature Med*., **2**, 689–692.
O'Reilly,M.S. *et al*. (1997) Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell*, **88**, 277–285.
Pan,T.C., Sasaki,T., Zhang,R.-Z., Fässler,R., Timpl,R. and Chu,M.-L. (1993) Structure and expression of fibulin-2, a novel extracellular matrix protein with multiple EGF-like repeats and consensus motifs for calcium-binding. *J. Cell Biol*., **123**, 1269–1277.
Rehn,M. and Pihlajaniemi,T. (1994) α1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen. *Proc. Natl Acad. Sci. USA*, **91**, 4234–4238.
Rehn,M., Hintikka,E. and Pihlajaniemi,T. (1994) Primary structure of the α1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the α1(XVIII) chain with its homologue, the α1(XV) collagen chain. *J. Biol. Chem*., **269**, 13929–13935.
Rehn,M., Hintikka,E. and Pihlajaniemi,T. (1996) Characterization of the mouse gene for the α1 chain of type XVIII collagen (Col 18a1) reveals that three variant N-terminal polypeptides forms are transcribed from two widely separated promoters. *Genomics*, **32**, 436–446.
Reinhardt,D.P., Sasaki,T., Dzamba,B.J., Keene,D.R., Chu,M.-L., Göhring,W., Timpl,R. and Sakai,L.Y. (1996) Fibrillin-1 and fibulin-2 interact and are colocalized in some tissues. *J. Biol. Chem.*, **271**, 19489–19496.
Sasaki,T., Göhring,W., Pan,T.-C., Chu,M.-L. and Timpl,R. (1995) Binding of mouse and human fibulin-2 to extracellular matrix ligands. *J. Mol. Biol*., **254**, 892–899.
Sasaki,T., Wiedemann,H., Matzner,M., Chu,M.-L. and Timpl,R. (1996) Expression of fibulin-2 by fibroblasts and deposition with fibronectin into a fibrillar matrix. *J. Cell Sci*., **109**, 2895–2904.
Sasaki,T., Hohenester,E., Göhring,W. and Timpl,R. (1998) Crystal structure and mapping by site-directed mutagenesis of the collagen-binding epitope of an activated form of BM-40/SPARC/osteonectin. *EMBO J*., **17**, 1625–1634.
Schulze,B., Mann,K., Battistutta,R., Wiedemann,H. and Timpl,R. (1995) Structural properties of recombinant domain III-3 of perlecan containing a globular domain inserted into an epidermal-growth-factor-like motif. *Eur. J. Biochem*., **231**, 551–556.
Ständker,L., Schrader,M., Kanse,S.M., Jürgens,M., Forssmann,W.-G. and Preissner,K.T. (1997) Isolation and characterization of the circulating form of human endostatin. *FEBS Lett*., **420**, 129–133.
Takahashi,K. (1968) The reaction of phenylglyoxal with arginine residues in proteins. *J. Biol. Chem*., **243**, 6171–6179.
Taraboletti,G., Roberts,D., Liotta,L.A. and Giavazzi,R. (1990a) Platelet thrombospondin modulates endothelial cell adhesion, mobility and growth: a potential angiogenesis regulatory factor. *J. Cell Biol*., **111**, 765–772.
Taraboletti,G., Rao,C.N., Krutzsch,H.C., Liotta,L.A. and Roberts,D.D. (1990b) Sulfatide-binding domain of the laminin A chain. *J. Biol. Chem*., **265**, 12253–12258.
Timpl,R. (1982) Antibodies to collagens and procollagens. *Methods Enzymol*., **82**, 472–498.
Tolsma,S.S., Volpert,O.V., Good,D.J., Frazier,W.A., Polverini,P.J. and Bouck,N. (1993) Peptides derived from two separate domains of the matrix protein thrombospondin-1 have anti-angiogenic activity. *J. Cell Biol*., **122**, 497–511.
Vogel,T., Guo,N., Krutzsch,H.C., Blake,D.A., Hartmann,J., Mendelovitz,S., Panet,A. and Roberts,D.D. (1993) Modulation of endothelial cell proliferation, adhesion and motility by recombinant heparin-binding domain and synthetic peptides from the type I repeats of thrombospondin. *J. Cell. Biochem*., **53**, 74–84.

Received May 21, 1998; revised and accepted June 11, 1998