FEBS Letters 411 (1997) 164–168 FEBS 18834

Sequence and expression of a novel member (LTBP-4) of the family of latent transforming growth factor-β binding proteins

Richard Giltay, Günter Kostka, Rupert Timpl*

Max-Planck-Institut für Biochemie, Am Klopferspitz 18A, D-82152 Martinsried, Germany

Received 22 May 1997

Abstract Overlapping cDNA clones from human heart and melanoma libraries were used to establish the 1587-residue sequence of a novel protein (LTBP-4) belonging to the family of extracellular microfibrillar proteins which also bind transforming growth factor-β. LTBP-4 consists of 20 EG modules, 17 of them with a consensus sequence for calcium binding, 4 TB modules with 8 cysteines and several proline-rich regions. Northern blots demonstrated a single 5 kb mRNA which is highly expressed in heart but also present in skeletal muscle, pancreas, placenta and lung. The modular structure predicts that LTBP-4 should be a microfibrillar protein which probably also binds TGF-β.

Key words: Extracellular matrix; Microfibrillar protein; TGF-β1 binding; Tissue expression

1. Introduction

A large and still increasing family of extracellular and transmembrane proteins is characterized by long tandem arrays of epidermal growth factor-like (EG) modules which possess a consensus sequence for calcium binding [1]. Typical members are the microfibrillar proteins fibrillin [2,3] and fibulin [4,5], each of which exists as two isoforms and basically has a rod-like shape. This shape is primarily determined by the EG modules which are stabilized by calcium binding as shown for an EG dimer of fibrillin-1 [6]. A further subfamily closely related to the fibrillins consists of the latent transforming growth factor-β binding proteins (LTBP) which were originally identified in a covalent complex with latent TGF-β1 in platelet extracts [7,8]. So far, complete sequences have been reported for three isoforms, LTBP-1 [9–11], LTBP-2 [12–14] and LTBP-3 [15]. Apart from a large and variable number of EG modules with six cysteines, they also contain four LTBP-like (TB) modules, characterized by an eight cysteine pattern [16], which they share with the fibrillins. Recent evidence indicates that they are associated with extracellular microfibrils containing fibronectin in fibroblast cultures [17], with elastic fibrillin-containing microfibrils [13], bone nodules [18] and microfibrillar structures that are important for epithelial-mesenchymal transitions in heart development [19]. LTBP-1 and -2 also bound to the extracellular matrix deposited by cultured fibroblasts, endothelial and epithelial cells [12,20,21]. This binding is apparently dependent on the N-terminal region of LTBP-1 [11,22]. Together, the data identify LTBPs as typical extracellular matrix proteins, although their supramolecular organisation still remains to be elucidated.

As well as their functions as matrix components, all three isoforms of LTBP share the unique property of a disulphide-mediated association with the latent form of TGF-β1 [7,8,12,15]. In this complex, mature TGF-β1 is already proteolytically released from the precursor form but still in a non-covalently associated and inactive state. Formation of the complex occurs fast and intracellularly and considerably enhances TGF-β1 secretion [23]. The mechanism of activation is still unclear but may include LTBP-mediated presentation to cell surfaces [24], possibly facilitated by the proteolytic release of LTBP from the matrix [11,21]. The binding of TGF-β1 to LTBP-1 was recently shown to occur through the TB-3 module located close to its C-terminal region and very likely involves a disulphide exchange with the precursor portion of TGF-β1 [22,25].

In a recent search for members belonging to this family of extracellular matrix proteins, we identified a particular expressed sequence tag (EST) cDNA clone as a further candidate gene. Complete cloning and sequence analysis demonstrated the existence of a novel LTBP-4 with qualifications to be a microfibrillar and probably additional TGF-β1 binding protein. Its tissue expression patterns indicated overlap with as well as differences from previously identified LTBP isoforms.

2. Materials and methods

A cDNA clone Z19101 from a human heart library (Stratagene; genebank accession number Z19101) was kindly supplied by Dr. B. Obermeier (Genzentrum, University Munich). After random primed labelling (Stratagene), it was used to screen a random primed human melanoma λ ZAP library (Stratagene) using standard hybridisation conditions (0.5 M sodium phosphate, pH 7.2, 7% SDS, 1 mM EDTA, 1% BSA) overnight at 65°C. Filters were then washed once in the same buffer (10 min, 65°C) and twice in 0.04 M sodium phosphate, pH 7.2, 1% SDS (15 min, 65°C) and exposed to X-ray film. Plating, isolation and purification of phage plaques were performed according to standard protocols. After excision of the inserts, the resulting clones were subcloned into the *E. coli* strain DH5α. cDNAs were isolated according to standard protocols and further analysed by Southern hybridisation. Subsequently, clones were sequenced with specific primers using a cycle sequencing terminator kit (Perkin Elmer) and an automated 373A DNA sequencer (Applied Biosystems). Deduced amino acid sequences were compared by the GCG program pile up and CLUSTAL W.

A commercially available Northern blot containing poly-A^+ RNA isolated from eight different human tissues (Clontech) was purchased and hybridized to an *Apa*I fragment (bp 3069–3597) of LTBP-4. The fragment was labelled using the Stratagene random primed labelling kit supplemented with α-[^{32}P]dCTP according to the manufacturer's instructions. Hybridisation was performed in 0.5 M sodium phosphate, pH 7.2, 7% SDS, 1% BSA, 1 mM EDTA, overnight at 65°C. Final washes were with 0.04 M sodium phosphate, pH 7.2, 1% SDS, 1 mM EDTA for 15 min at 65°C. After washing the blot was air dried and exposed to X-ray film.

*Corresponding author. Fax: (+44) 89-8578-2422.

PII S0014-5793(97)00685-6

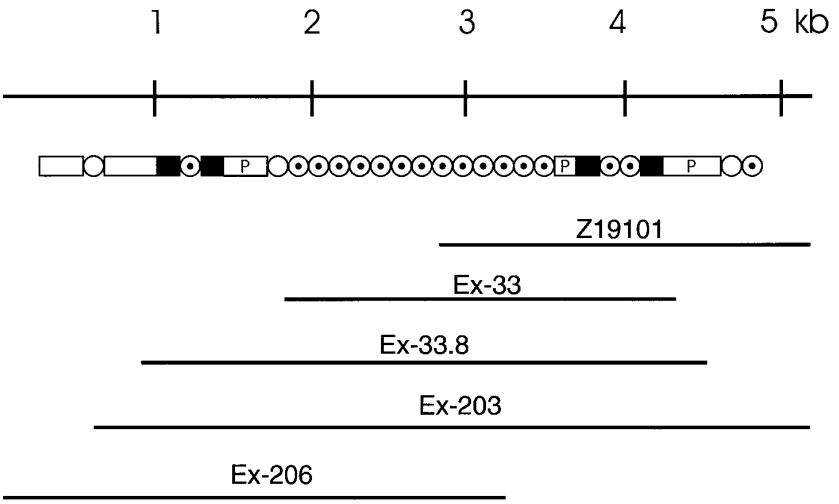


Fig. 1. Scheme of the module arrangement of human LTBP-4 and correlation with cDNA clones used for sequence analysis. The modules include EG (circles), EG with calcium-binding sequence (circles with dot) and TB (black bars). Open bars indicate non-classified sequences and are marked with P if rich in proline.

3. Results

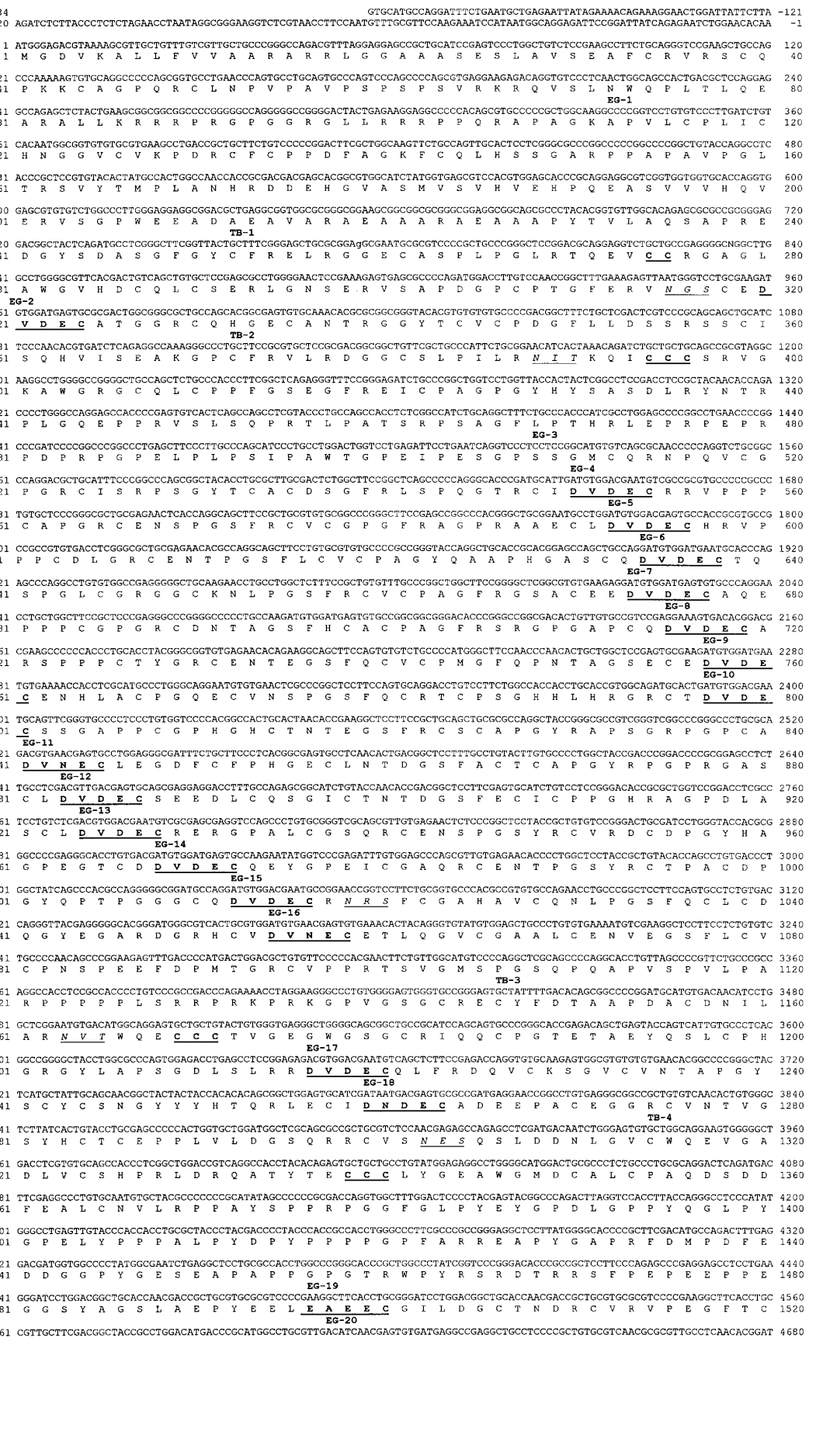
The computer-assisted screening of an EST library (Merck/WashU project) with the sequence of the EG repeats of fibulin-1 [26] identified a related 2 kb cDNA clone Z19101. Since sequence analysis of this clone revealed tandem arrays of calcium-binding EG modules, we extended the screening to a human melanoma λ Zap library and identified four more overlapping clones spanning a size of slightly more than 5 kb (Fig. 1). Their complete sequence analysis (Fig. 2) demonstrated an open reading frame of 4761 bp joined to 184 bp untranslated region at the 5′-end and 267 bp at the 3′-end with a typical polyadenylation signal starting at position 4996. 32 bp after this position is a long poly-A tail.

The start of the amino acid sequence (1587 residues) was assigned to a methionine preceding a typical hydrophobic signal peptide sequence. Predictions as to its cleavage site [27] indicated position 15/16 or as a second best choice position 23/24. The postulated start methionine is preceded by many stop codons in all three reading frames and therefore very likely correctly assigned. The deduced amino acid sequence could not be found in the databank and predicts a polypeptide with a calculated molecular mass of 187 kDa and a pI of 5.87. Five putative N-glycosylation sites could be identified at positions 315, 388, 1018, 1163 and 1302 (Fig. 2) and, if occupied, may increase the molecular mass. The sequence is dominated by a total of 24 cysteine-rich motifs which could be clearly identified as either EG or TB modules [16]. Their arrangement is most similar to those shown for other LTBP isoforms [9–15], hence the novel polypeptide will be referred to as LTBP-4.

The hallmark of LTBP-4 structure is a central array of 14 consecutive EG modules (EG-3 to EG-16), which, apart from the most N-terminal one, all possess a consensus sequence for calcium binding (Figs. 1 and 2). This consensus structure is characterized by an almost invariable DVDE sequence in front of the first cysteine which provides some major calcium ligation sites [28]. The C-terminal region is characterized by a tandem array TB-3/EG-17,18/TB-4 and directly at the C-terminus by the last pair of EG modules. The N-terminal region contains another TB-1/EG-2/TB-2 array and a single EG-1 module. From the total of 20 EG modules, 17 can be predicted to bind calcium. These regions of defined protein motifs are connected or terminated by non-classified segments of 50–130 residues some of which are rich in proline (26–30%). Peculiar features include a short segment with four cysteines close to the N-terminus and one odd cysteine (position 318) in front of module EG-2. LTBP-4 did not contain a potentially cell-adhesive RGD sequence.

The other three LTBP isoforms show basically the same organisation of EG and TB modules into several clusters, but have a smaller number of EG modules (8–13) in the central array. A comparison of the amino acid sequences of these clusters showed identities in the range of 42–59% (Table 1). The overall sequence identity was somewhat lower (38–42%). Certain sequence peculiarities, such as a missing cysteine in the CCC sequence of the TB-1 module (Fig. 2), are, however, maintained in all four isoforms.

Further differences between LTBP-4 and the other isoforms were observed by examining its expression by Northern blotting of poly-A^+ RNA from eight human tissues (Fig. 3). A cDNA probe with only 60–66% nucleotide sequence identity to LTBP-1, -2 and -3 was used under stringent hybridisation conditions in order to eliminate cross-hybridisation and the commercially available mRNA filter was similar to those showing distinct and different patterns for human LTBP-1 and LTBP-2 [12]. The LTBP-4 mRNA appeared as a single band of about 5 kb, in agreement with the size of the cDNA sequence. A particular strong expression was observed in heart followed by skeletal muscle and pancreas. Moderate signals were observed for placenta and lung and no or only weak ones for brain, liver and kidney. Northern blots for human LTBP-1 were previously shown to hybridize to two mRNAs of 5.2 and 7 kb and for human LTBP-2 to bands of 7.5 and 9.5 kb [11–13]. None of these bands were detected with the cDNA probe for human LTBP-4, demonstrating the specificity of hybridisation. Further preliminary in situ hybridisation experiments in adult mouse heart demonstrated LTBP-4 expression in interstitial fibroblasts of heart valves but not in the myocardium (unpublished). A restricted expression in



```
1521 R C F D G Y R L D M T R M A C V D I N E C D E A E A A S P L C V N A R C L N T D 1560

4681 GGCTCCTTCCGCTGCATCTGCCGCCCGGGATTCGCACCCACGCACCAGCCGCACCACTGTGCGCCCGCACGGCCCCGGGCCTGAGCCCTGGCACCCGCTGGCCACCCACCCGCGCCCGCC 4800
1561 G S F R C I C R P G F A P T H Q P H H C A P A R P R A *                                                                  1587

4801 ACTCGGGGCCCCTGCCGCGCATCCTGCAGCCCGCTTATGCGTATGTGCACGGGGCCGCCCCGCCTGGACCTGGAGAAGGGACCTACGGACGCCTGGAAGCTGCGACGCCCTGCACTGCTCC 4920
4921 CGCCTCCACCAGCGCCTCCCACTGATGTCGTGGTCCCGGGCCTGGCCCAGGGGCCCCTTTACATGCCCTCTCCCTTTTATAAAATTTTCCATTAAAAACCACCTATTTTCTAAAAAAAAA 5040
```

Fig. 2. Complete cDNA and deduced amino acid sequence of human LTBP-4. The start of individual EG (nos. 1–20) and TB modules (nos. 1–4) are denoted on top of the nucleotide sequence (EMBL Nucleotide Sequence Database, accesion no. Y13622). Highly conserved sequences of a calcium-binding consensus motif (EG) and a triple cysteine (TB) are shown in bold letters and are underlined. Five potential N-glycosylation sites are underlined and shown in italics. Potential signal peptide cleavage sites include positions 15/16 and 23/24. A potential polyadenylation signal is shown in bold and precedes a long poly-A tail.

certain tissue compartments has also been observed for mouse LTBP-3 [15].

4. Discussion

A distinct family of extracellular matrix proteins referred to as LTBP-1 to LTBP-3 were classified together as a subfamily of microfibrillar proteins based on their modular structure and the ability to associate with latent TGF-β1 in a covalent manner [9–15]. We have now identified the cDNA sequence of a fourth member of this subfamily, LTBP-4, which has a very similar modular structure. More distantly related proteins include the fibrillins [29], which share the calcium-binding EG and TB modules, and the fibulins [26,30], in which the TB modules are replaced by other sets of modules. Proteins belonging to the latter two subfamilies have basically a rod-like structure [2–5]. This has not yet been demonstrated for LTBPs, but their association with microfibrils in tissues and cultured cells [13,17–19] strongly favours this possibility. Such fibrillar structures are apparently dependent on calcium-binding EG modules, as shown in a recent study at atomic resolution [6]. Calcium binding also enhances protease stability of LTBP-1 [31], fibulins [32] and fibrillin [33].

The reason for the existence of several LTBP isoforms is still unclear, reflecting the fact that, except for TGF-β1 binding, they have been insufficiently characterized at the functional level. The different, although overlapping, tissue expression of LTBP mRNAs [12] would indicate functional differences. LTBP-4 is not only strongly expressed in heart like LTBP-1, but also in skeletal muscle and pancreas. Expression of LTBP-1 and LTBP-2 could not be demonstrated for pancreas [12]. These analyses showed two splice variants which differ by more than 1 kb for LTBP-1 [11] and probably for LTBP-2 mRNA [12,13]. Substantial splicing has not yet been detected for LTBP-3 [15] and, as described here, was not found for LTBP-4. Splice variation has been localized to the N-terminal region of LTBP-1 and may determine the strength of its association with the extracellular matrix [11,22]. A particularly high activity was found for the long form of LTBP-1, which contains an additional EG module not present in LTBP-3 and LTBP-4. Human LTBP-1 [10] and LTBP-2 [12] but not LTBP-4 contain a single RGD sequence and may therefore differ in integrin-mediated cell adhesiveness.

Based on the sequence data, LTBP-4 is very likely a rod-like protein and associated with microfibrillar tissue structures. Its covalent association with latent TGF-β1 remains an open question. This association was shown to occur through module TB-3 in recombinant LTBP-1 fragments and involves a disulphide exchange with latent TGF-β1 [22,25]. A sequence comparison of the TB-3 modules of all four LTBP isoforms shows a high conservation in certain regions close to the three cysteine sequence, which could indicate conservation of the binding structure (Fig. 4). The binding epitope has not yet been precisely mapped but does not require glycosylation of an invariant NVT sequence [22]. The availability of cDNA clones for LTBP-4 will now allow its binding to TGF-β1 and its molecular shape to be examined by recombinant production. Such approaches have already been successful for binding fragments of fibrillin [34] and fibulins [5].

Acknowledgements: We are grateful to Mrs. Christa Hintz-Martini and Mr. Marius Boicu for expert technical assistance and to Dr. B. Obermeier for a gift of cDNA and would like to thank the Deutsche Forschungsgemeinschaft (SFB 266) for financial support.

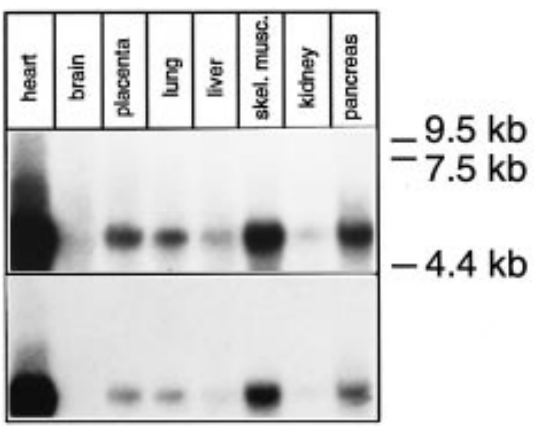


Fig. 3. Expression of human LTBP-4 in different tissues examined by Northern blots. The same filter was exposed for different periods (top: 8 h; bottom: 2 h) in order to demonstrate the expression of a single 5 kb mRNA which existed in variable amounts.

Table 1
Sequence identities of LTBP-4 with other LTBP isoforms

LTBP-4 structure compared	Sequence identity (%) to		
	LTBP-1	LTBP-2	LTBP-3
EG-1	56	47	59
TB-1/EG-2/TB-2	47	50	44
EG-3 to EG-16[a]	48	49	46
TB-3/EG-17,18/TB-4	50	46	42
EG-19,20	59	57	47
Entire protein	42	42	38

Individual module segments or the total sequence are compared with human LTBP-1 [10], human LTBP-2 [12,14] and mouse LTBP-3 [15].
[a]Adjusted to the best score for the smallest number of shared modules.

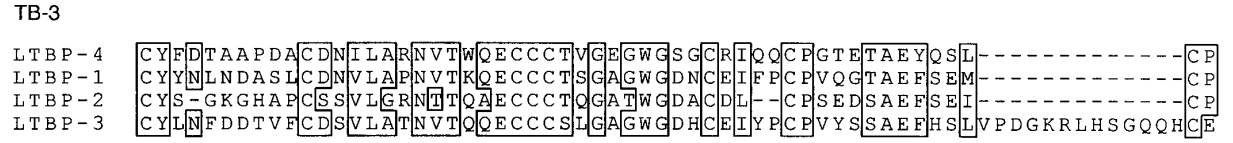


Fig. 4. Sequence alignment of the TB-3 module of all four LTBP isoforms. This module is involved in TGF-β1 binding in LTBP-1. Identical or very similar residues that occur in each sequence are boxed.

References

[1] Campbell, I.D. and Bork, P. (1993) Curr. Opin. Struct. Biol. 3, 385–392.
[2] Sakai, L.Y., Keene, D.R., Glanville, R.W. and Bächinger, H.P. (1991) J. Biol. Chem. 266, 14763–14770.
[3] Reinhardt, D.P., Keene, D.R., Corson, G.M., Pöschl, E., Bächinger, H.P., Gambee, J.E. and Sakai, L.Y. (1996) J. Mol. Biol. 258, 104–116.
[4] Sasaki, T., Wiedemann, H., Matzner, M., Chu, M.-L. and Timpl, R. (1996) J. Cell Sci. 109, 2895–2904.
[5] Sasaki, T., Mann, K., Wiedemann, H., Göhring, W., Lustig, A., Engel, J., Chu, M.-L. and Timpl, R. (1997) EMBO J., in press.
[6] Downing, A.K., Knott, V., Werner, J.M., Cardy, C.M., Campbell, I.D. and Handford, P.A. (1996) Cell 85, 597–605.
[7] Miyazono, K., Hellman, U., Wernstedt, C. and Heldin, C.-H. (1988) J. Biol. Chem. 263, 6407–6415.
[8] Wakefield, L.M., Smith, D.M., Flanders, K.C. and Sporn, M.B. (1988) J. Biol. Chem. 263, 7646–7654.
[9] Tsuji, T., Okada, F., Yamaguchi, K. and Nakamura, T. (1990) Proc. Natl. Acad. Sci. USA 87, 8835–8839.
[10] Kanzaki, T., Olofsson, A., Morén, A., Wernstedt, C., Hellman, U., Miyazono, K., Claesson-Welsh, L. and Heldin, C.-H. (1990) Cell 61, 1051–1061.
[11] Olofsson, A., Ichijo, H., Morén, A., ten Dijke, P., Miyazono, K. and Heldin, C.-H. (1995) J. Biol. Chem. 270, 31294–31297.
[12] Morén, A., Olofsson, A., Stenman, G., Sahlin, P., Kanzaki, T., Claesson-Welsh, L., ten Dijke, P., Miyazono, K. and Heldin, C.-H. (1994) J. Biol. Chem. 269, 32469–32478.
[13] Gibson, M.A., Hatzinikolas, G., Davis, E.C., Baker, E., Sutherland, G.R. and Mecham, R.P. (1995) Mol. Cell. Biol. 15, 6932–6942.
[14] Bashir, M.M., Han, M.-D., Abrams, W.R., Tucker, T., Ma, R.-I., Gibson, M., Ritty, T., Mecham, R. and Rosenbloom, J. (1996) Int. J. Biochem. Cell Biol. 28, 531–542.
[15] Yin, W., Smiley, E., Germiller, J., Mecham, R.P., Florer, J.B., Wenstrup, R.J. and Bonadio, J. (1995) J. Biol. Chem. 270, 10147–10160.
[16] Bork, P., Downing, A.K., Kieffer, B. and Campbell, I.D. (1996) Quart. Rev. Biophys. 29, 119–167.
[17] Taipale, J., Saharinen, J., Hedman, K. and Keski-Oja, J. (1996) J. Histochem. Cytochem. 44, 875–889.
[18] Dallas, S.L., Miyazono, K., Skerry, T.M., Mundy, G.R. and Bonewald, L.F. (1995) J. Cell Biol. 131, 539–549.
[19] Nakajima, Y., Miyazono, K., Kato, M., Takase, M., Yamagishi, T. and Nakamura, H. (1997) J. Cell Biol. 136, 193–204.
[20] Taipale, J., Miyazono, K., Heldin, C.-H. and Keski-Oja, J. (1994) J. Cell Biol. 124, 171–181.
[21] Taipale, J., Lohi, J., Saarinen, J., Kovanen, P.T. and Keski-Oja, J. (1995) J. Biol. Chem. 270, 4689–4696.
[22] Saharinen, J., Taipale, J. and Keski-Oja, J. (1996) EMBO J. 15, 245–253.
[23] Miyazono, K., Olofsson, A., Colosetti, P. and Heldin, C.-H. (1991) EMBO J. 10, 1091–1101.
[24] Flaumenhaft, R., Abe, M., Sato, K., Miyazono, K., Harpel, J., Heldin, C.-H. and Rifkin, D.B. (1993) J. Cell Biol. 120, 995–1002.
[25] Gleizes, P.E., Beavis, R.C., Mazzieri, R., Shen, B. and Rifkin, D.B. (1996) J. Biol. Chem. 271, 29891–29896.
[26] Pan, T.-C., Kluge, M., Zhang, R.-Z., Mayer, U., Timpl, R. and Chu, M.-L. (1993) Eur. J. Biochem. 215, 733–740.
[27] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[28] Selander-Sunnerhagen, M., Ullner, M., Persson, E., Teleman, O., Stenflo, J. and Drakenberg, T. (1992) J. Biol. Chem. 267, 19642–19649.
[29] Kielty, C.M. and Shuttleworth, C.A. (1995) Int. J. Biochem. Cell Biol. 27, 747–760.
[30] Pan, T.-C., Sasaki, T., Zhang, R.-Z., Fässler, R., Timpl, R. and Chu, M.-L. (1993) J. Cell Biol. 123, 1269–1277.
[31] Colosetti, P., Hellman, U., Heldin, C.-H. and Miyazono, K. (1993) FEBS Lett. 320, 140–144.
[32] Sasaki, T., Mann, K., Murphy, G., Chu, M.-L. and Timpl, R. (1996) Eur. J. Biochem. 240, 427–434.
[33] Reinhardt, D.P., Ono, R.N. and Sakai, L.Y. (1997) J. Biol. Chem. 272, 1231–1236.
[34] Reinhardt, D.P., Sasaki, T., Dzamba, B.J., Keene, D.R., Chu, M.-L., Göhring, W., Timpl, R. and Sakai, L.Y. (1996) J. Biol. Chem. 271, 19489–19496.