



Identification of binding partners interacting with the $\alpha 1$ -N-propeptide of type V collagen

Sofie Symoens, Marjolijn Renard, Christelle Bonod Bidaud, Delfien Syx, Elisabeth Vaganay, Fransiska Malfait, Sylvie Ricard-Blum, Efrat Kessler, Lut van Laer, Paul Coucke, et al.

► To cite this version:

Sofie Symoens, Marjolijn Renard, Christelle Bonod Bidaud, Delfien Syx, Elisabeth Vaganay, et al.. Identification of binding partners interacting with the $\alpha 1$ -N-propeptide of type V collagen. Biochemical Journal, 2010, 433 (2), pp.371-381. 10.1042/BJ20101061 . hal-00549896

HAL Id: hal-00549896

<https://hal.science/hal-00549896>

Submitted on 23 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins**IDENTIFICATION OF BINDING PARTNERS INTERACTING WITH THE $\alpha 1$ -N-PROPEPTIDE OF TYPE V COLLAGEN**

Sofie Symoens^{*}, Marjolijn Renard^{*}, Christelle Bonod-Bidaud[†], Delfien Syx^{*}, Elisabeth Vaganay[†], Fransiska Malfait^{*}, Sylvie Ricard-Blum[†], Efrat Kessler[‡], Lut Van Laer^{*}, Paul Coucke^{*}, Florence Ruggiero[†], Anne De Paepe^{*,||}

From the ^{}Center for Medical Genetics, Ghent University Hospital, 9000 Ghent, Belgium, the [†]Institut de Biologie et Chimie des Protéines, CNRS UMR 5086, Université Lyon1, Université de Lyon, IFR128 Biosciences-Gerland, France, and the [‡]Maurice and Gabriela Goldschleger Eye Research Institute, Tel Aviv University Sackler Faculty of Medicine, Sheba Medical Center, Tel-Hashomer 52621, Israel*

^{||}To whom correspondence should be addressed. Center for Medical Genetics, Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium. Tel: +32-9-332.36.03. Fax: +32-9-332.49.70. E-mail: Anne.DePaepe@UGent.be

Short title: Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

Key words: Ehlers-Danlos syndrome / Type V collagen / Yeast-two-hybrid / Surface Plasmon resonance / co-immunoprecipitation / interaction partners

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins**Abstract**

The predominant form of type V collagen is the $[\alpha 1(V)]_2\alpha 2(V)$ heterotrimer. Mutations in *COL5A1* or *COL5A2*, encoding respectively the $\alpha 1(V)$ - and $\alpha 2(V)$ -collagen chain, cause classic Ehlers-Danlos syndrome (EDS), a heritable connective tissue disorder, characterized by fragile hyperextensible skin and joint hypermobility. Approximately half of the classic EDS cases remains unexplained.

Type V collagen controls collagen fibrillogenesis through its conserved $\alpha 1(V)$ -amino(N)-propeptide domain. To gain insight into the role of this domain, a yeast-two-hybrid screen among proteins expressed in human dermal fibroblasts was performed utilizing the amino(N)-propeptide as a bait. We identified 12 interacting proteins including extracellular matrix proteins and proteins involved in collagen biosynthesis. Eleven interactions were confirmed by surface plasmon resonance and/or co-immunoprecipitation: $\alpha 1(I)$ - and $\alpha 2(I)$ -collagen chains, $\alpha 1(VI)$ -, $\alpha 2(VI)$ - and $\alpha 3(VI)$ -collagen chains, tenascin-C, fibronectin, procollagen C-proteinase enhancer-1 (PCPE-1), tissue inhibitor of metalloproteinases-1, matrix metalloproteinase-2 and transforming growth factor- $\beta 1$. Solid-phase binding assays confirmed the involvement of the $\alpha 1(V)$ -N-propeptide in the interaction between native type V collagen and type VI collagen, suggesting a bridging function of this protein complex in the cell-matrix environment. Enzymatic studies showed that processing of the $\alpha 1(V)$ -N-propeptide by the procollagen C-proteinase BMP-1 is enhanced by PCPE-1. These interactions are likely involved in ECM homeostasis and their disruption could explain the pathogenetic mechanism in unresolved classic EDS cases.

Abbreviations used: pN $\alpha 1(V)$, construct encompassing the $\alpha 1(V)$ -N-propeptide and part of the type V collagen triple helix; BMP-1, procollagen C-proteinase, also known as procollagen C-proteinase; COL1, major triple helix domain; COL2, short interrupted triple helix domain; ECM, extracellular matrix; EDS, Ehlers-Danlos syndrome; NC2, flexible hinge region; NC3, non-collagenous 3 domain (TSPN-1 and VAR); PARP, proline/arginine-rich protein; RU, resonance unit; SPR, surface plasmon resonance; TSPN-1, thrombospondin-1 N-terminal domain like; VAR, variable region; Y2H, yeast-2-hybrid

Introduction

The extracellular matrix (ECM) is composed of a complex combination of macromolecules organized into tissue-specific networks. The major structural ECM components are the fibrillar collagens, which aggregate into highly organized banded collagen fibrils. The low-abundant fibrillar type V collagen plays an important role in the control of type I/V collagen fibrillogenesis, as illustrated by studies in genetically modified mouse models, which demonstrated that type V collagen is involved in fibril nucleation and regulation of collagen fibril diameter [1-3]. Furthermore, although embryonic fibroblasts of *col5a1*^{-/-} mice synthesize and secrete normal amounts of type I collagen, collagen fibrils are virtually absent and mice die at the onset of organogenesis [3]. These findings indicate that a certain threshold of type V collagen should be maintained for normal collagen fibrillogenesis.

The most broadly distributed isoform of type V collagen is the $[\alpha 1(V)]_2\alpha 2(V)$ heterotrimer [1]. During fibril formation, this heterotrimer is incorporated into type I collagen fibrils, thereby regulating the diameter of these fibrils through partial retention of the $\alpha 1(V)$ -N-propeptide [1, 3]. The current working model predicts that the entire type V collagen triple helix is buried within the fibril, whereas type I collagen is present along the fibril surface. The $\alpha 1(V)$ -N-propeptide extends outwards from the gap zones of the heterotypic fibrils and is thus the only part of the type V collagen molecule that emerges at the surface of these fibrils [1].

The $\alpha 1(V)$ -N-propeptide exhibits a unique multidomain structure (Figure 1A): a TSPN-1 region (thrombospondin-1 N-terminal domain like), also known as the PARP (Proline/Arginine-Rich

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

Protein) domain, followed by a variable region (VAR), which together form the non-collagenous 3 domain (NC3 domain) and a short interrupted triple helix (COL2). The COL2 region is adjacent to a short stretch, the NC2 domain. This flexible region, that precedes the $\alpha 1(V)$ major triple helix (COL1), allows the N-propeptide to project away from the major axis of the molecule. As such, the COL2 domain extends across the gap zone, placing the non-collagenous domains at the surface of heterotypic type I/V collagen fibrils [1]. The $\alpha 2(V)$ -N-propeptide is shorter and less complex than its $\alpha 1(V)$ counterpart. It is composed of a cysteine-rich (CR) domain, which precedes a small COL2 domain (Figure 1A) [4]. Processing of the $\alpha 1(V)$ -N-propeptide is complex as it occurs in a tissue-specific manner at different sites and can be mediated by BMP-1 [4, 5] and/or ADAMTS-2 enzymes [6] (Figure 1), whereas the $\alpha 2(V)$ -N-propeptide remains intact in the mature molecule [4]. While BMP-1 cleavage of the $\alpha 1(V)$ -N-propeptide - which represents the *in vivo* dominant processing activity- releases the TSPN-1 domain and retains the VAR, COL2 and NC2 domains [4] (Figure 1A and B), ADAMTS-2 cleavage releases the entire NC3 domain (TSPN-1 and VAR) (Figure 1A) [6].

Dysregulation of collagen fibrillogenesis is a hallmark of several subtypes of the Ehlers-Danlos syndrome (EDS). Classic EDS (MIM# 130000, MIM# 130010) is characterized by fragile and hyperextensible skin, poor wound healing with atrophic scarring and joint hypermobility. As a result of impaired collagen fibrillogenesis, the dermis of classic EDS patients shows a decrease in fibril density and the presence of large irregular collagen fibrils (called collagen 'cauliflowers') [7]. Approximately 50% of the patients with classic EDS harbour a heterozygous mutation in *COL5A1* or *COL5A2*, encoding the $\alpha 1(V)$ - and the $\alpha 2(V)$ -collagen chains respectively. Most commonly, mutations introduce a premature termination codon, thereby leading to a non-functional *COL5A1* allele and a decreased amount of type V collagen protein (*COL5A1* haploinsufficiency) [8, 9]. A minority of mutations affects the type V collagen structure, including a number of glycine substitutions and exon skipping mutations in the collagen triple helix, as well as C-propeptide mutations which interfere with the incorporation of the mutant chains into the type V collagen heterotrimer [8, 9]. Recently, we have shown that mutations in the $\alpha 1(V)$ -signal peptide result in decreased type V collagen secretion into the ECM [10]. Interestingly, until now, only two mutations in the highly conserved $\alpha 1(V)$ -N-propeptide have been described allowing the expression of the mutant protein: a single homozygous amino acid substitution (p.G530S) [11] and a splice-acceptor mutation (IVS4-2A>G) [12] (Figure 1A). While the p.G530S substitution is located in the COL2 domain and is suggested to interfere with the correct folding of this region, the IVS4-2A>G splice-acceptor mutation generates different mutant transcripts, affecting the VAR region and deleting the BMP-1 cleavage site and two of the four cysteines residues. These latter transcripts are translated and mutant proteins with abnormal $\alpha 1(V)$ -N-propeptides are secreted and interfere with normal collagen fibrillogenesis.

Various binding partners to type V collagen have already been identified in earlier studies, including thrombospondin, insulin, the small leucine-rich proteoglycan decorin, heparin (as reviewed in Fichard et al., 1995 [13]), heparan sulphate [14], biglycan [15], osteonectin/SPARC [16], tenascin-X [17], integrins [18], discoidin domain receptors 1 and 2 (DDR1 and DDR2) [19], uPARAP (urokinase Plasminogen Activator Receptor-Associated Protein) [20]. However, since pepsinized type V collagen was used for all these studies, the identified ligands only bind to the type V collagen triple helix, as the N- and the C-propeptides are removed during pepsin digestion. In this study, we intended to identify proteins that interact with the $\alpha 1(V)$ -N-propeptide domain using the yeast-two-hybrid (Y2H) approach. We selected the $\alpha 1(V)$ -N-propeptide because this is the only domain that can mediate type V collagen interactions in the extracellular matrix as it protrudes from the surface of heterotypic type I/V collagen fibrils. Furthermore, we decided to generate a human dermal skin cDNA library as a prey library because the skin is the major affected tissue in classic EDS patients. Identifying new interaction partners of the $\alpha 1(V)$ -N-propeptide will yield further insights into the process of collagen assembly and the

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

pathogenesis of (classic) EDS. Moreover, these molecules may represent new candidate genes for hitherto unresolved EDS cases.

Methods

Preparation of bait construct, library preparation and Y2H experiment

To screen for new interactions between the $\alpha 1(V)$ -N-propeptide and proteins expressed in the ECM of human dermal fibroblasts, we used the Matchmaker Two-Hybrid System (Clontech Laboratories). All procedures were performed according to the manufacturer's protocol. In order to generate the Y2H bait construct, pGBKT7-pN $\alpha 1(V)$, we started from the previously described plasmid pCEP4-pN $\alpha 1(V)$ [5]. We performed a polymerase chain reaction (PCR) on the pCEP4-pN $\alpha 1(V)$ plasmid using primers (Supplementary Table 1) that introduced unique restriction endonuclease sites for unidirectional cloning into the yeast DNA binding domain (BD) vector pGBKT7 (Clontech Laboratories). The pGBKT7 vector carries the tryptophan reporter gene. To verify whether the pN $\alpha 1(V)$ insert was cloned in-frame, the resulting pGBKT7-pN $\alpha 1(V)$ bait vector was sequenced.

To generate our prey cDNA library, polyA⁺ RNA was isolated from human dermal fibroblasts and converted to cDNA with random primers (Clontech Laboratories). Competent yeast cells (strain AH109) were co-transformed with this dermal fibroblast library, the pGADT7-Rec prey vector (which contains the linearized activation domain (AD) and carries the leucine reporter gene (Clontech Laboratories)) and the pGBKT7-pN $\alpha 1(V)$ bait plasmid. Transformed yeast cells were spread on series of agar-containing plates deficient in (tryptophan), (leucine), (tryptophan and leucine), (tryptophan, leucine and histidine), (tryptophan, leucine, histidine and adenine)(quadruple dropout, QDO) and incubated for 4-6 days at 30°C. Yeast colonies in which a bait-prey interaction occurred, survived stringent nutritional selection. Surviving colonies were plated onto QDO plates containing the chromogenic substrate X- α -Gal (Clontech Laboratories). Yeast colonies turned blue when a bait-prey interaction occurred. Transformations were also carried out with positive control (pGADT7-RecT + pGBKT7-53) and negative control (pGADT7-RecT + pGBKT7-Lam) plasmids (Clontech Laboratories).

Selection and analysis of positive interactions

Blue colonies were transferred to fresh QDO X- α -Gal plates and incubated for 4-6 days at 30°C. This process was repeated twice and colonies that maintained their phenotype and survived stringent nutritional selection were considered positive clones. To screen the inserts (encoding possible interaction partners) of the pGADT7-Rec prey vector, we performed a PCR reaction (touch-down program 72,1°C - 65,1°C) on the surviving colonies using the 5'AD-LD-Amplimer and 3'AD-LD-Amplimer primers (Supplementary Table 1). The resulting PCR fragments were sequenced (ABI3730XL, Applied Biosystems), and a BLAST search was performed (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). We only retained sequences which contained an open reading frame and encoded proteins that were present in the GenBank database (<http://www.ncbi.nlm.nih.gov/Genbank/>). Sequences that were omitted included those that were out-of-frame, that corresponded to unknown genes, that started with a poly-A tail indicating that the gene was probably inverted and that contained multiple stop codons or open reading frames that could not be identified in the GenBank database. In addition, genes which were on a list of known false positives, that are frequently found in Y2H screens independently of the target used, were omitted from further analysis [21] (<http://www.fccc.edu:80/research/labs/golemis/InteractionTrapInWork.html>). Putative interacting proteins were retested by cotransformation in yeast AH109 cells and stringent nutritional selection.

Tissue extracts and recombinant proteins

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

To confirm the identified protein-protein interactions, tissue extracts and several native and recombinant proteins were used. Details are provided in supplementary data (details and Supplementary Figure 1). The purity of the proteins was analyzed by 6% or 10% SDS-PAGE. Gels were stained with Coomassie Blue.

Surface plasmon resonance (SPR) binding assays

Measurements were performed using a Biacore 3000 or a Biacore T100 (GE Healthcare). PCPE-1 [22], and the pN $\alpha 1(V)$ -protein were covalently immobilized to the dextran matrix of a CM5 sensor chip via their primary amine groups. The carboxymethylated dextran surface was activated by the injection of a mixture of 0.2M N-ethyl-N'-(diethylaminopropyl)carbodiimide (EDC) and 0.05M N-hydroxysuccinimide (NHS). PCPE-1 was diluted in 10mM maleate buffer pH6.0, and the pN $\alpha 1(V)$ -protein in acetic acid pH3.5 before injection over the activated surface of the sensor chip. Residual active groups were blocked by injection of 1M ethanolamine pH8.5. Immobilization levels of 1438 resonance units (RU) for PCPE-1, and 7318 RU for the pN $\alpha 1(V)$ -protein were obtained. A control flow cell was activated by the NHS-EDC mixture and deactivated by 1M ethanolamine pH8.5 without any coupled protein. Control sensorgrams were subtracted on-line from the sensorgrams obtained on immobilized PCPE-1 and pN $\alpha 1(V)$ -protein in order to yield specific binding responses. The $\alpha 1(V)$ -N-propeptide (65 μ g/ml) was injected at 20 μ l/min for 4 min over immobilized PCPE-1 with 50mM Tris buffer pH7.5 containing 0.15M NaCl and 0.005% (v/v) P20 surfactant as running buffer in the presence or absence of 5mM CaCl₂ [22]. TGF- $\beta 1$ (3.1 μ g/ml) was injected over the immobilized pN $\alpha 1(V)$ -protein at 30 μ l/min for 2 min in 10mM Hepes pH7.4 containing 0.15M NaCl and 0.005% (v/v) P20 surfactant. The sensor chip surface was regenerated either with a pulse of 2M NaCl and/or a pulse of 1.5M NaCl and 1.5M guanidinium chloride.

SPR arrays were handled in a Biacore Flexchip system (GE Healthcare), a commercially available high-density array platform, and analyzed as described previously [23]. The pN $\alpha 1(V)$ -protein was printed directly in triplicate at concentrations varying from 50 to 200 μ g/ml onto the gold surface of a Gold Affinity chip (GE Healthcare) using a non-contact PiezoArray spotter (Perkin Elmer). Six drops of 330pl each were delivered to the surface of the chip (total spotted volume: 1.9nl, spot diameter: 250-300 μ m, spotted amount: 100-400pg/spot). The chip was blocked with blocking buffer containing mammalian proteins (Superblock, Pierce) for 5x5 min. The blocked chip was then equilibrated with PBS - 0.05% Tween 20 (P3563, Sigma) at 500 μ l/min for 90 min. Human type I collagen (150nM), type VI collagen (250nM), and laminin 1 (used as a negative control) were diluted in PBS - 0.05% Tween 20, and flowed at 500 μ l/min over the chip surface for 25 min. The dissociation of the complexes was monitored in PBS - 0.05% Tween 20 for 40 min. Data collected from reference spots (gold surface) and buffer spots were subtracted from those collected on spotted proteins to obtain specific binding curves.

Co-immunoprecipitation assay

Protein-protein interactions between the pN $\alpha 1(V)$ -protein and respectively fibronectin (FN), $\alpha 1(VI)$ -collagen chain, tissue inhibitor of matrix metalloproteinases-1 (TIMP-1), matrix metalloproteinase-2 (MMP-2), tenascin-C (TNC), TGF- $\beta 1$ and the type I collagen C-propeptide were confirmed using the Dynabeads[®] Protein A System (Invitrogen). HEK293T cells expressing the pN $\alpha 1(V)$ -protein [5] were cultured in growth medium containing hygromycin (Invitrogen). The medium fraction was collected, concentrated using the Centrprep Centrifugal Filter Unit with Ultracel-30 membrane (Millipore) and stored until use at -20°C.

To facilitate the detection of the different interacting proteins by Western blot, a flag-tag was included at their N-terminus. Expression constructs encoding the interacting domains of TNC and MMP-2 were generated in the flag-tag vector pMET7-fC-SH2, while expression constructs encoding the interacting domains of the $\alpha 1(VI)$ -collagen chain, TIMP-1 and FN were generated in the flag-tag vector pMG2-p85 using primers enlisted in Supplementary Table 1. To investigate

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

whether the inserts of all expression plasmids were in-frame, each insert was sequenced (ABI3730XL, Applied Biosystems). Subsequently, HEK293T cells were transfected with the expression plasmids encoding the potential interacting domains of the $\alpha 1(V)$ -collagen chain, FN and TIMP1 using lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol. The HEK293T cells were transfected with the expression plasmids encoding the potential MMP-2 and TNC interacting domains using the CaCl_2 methodology. Transfected cells were lysed using M-PER buffer (Pierce), supplemented with Complete Protease Inhibitor Cocktail (Roche).

For the co-immunoprecipitation assay, the monoclonal anti- $\alpha 1(V)$ -N-propeptide antibody (18G5, directed against the NC2 domain of the $\alpha 1(V)$ -N-propeptide [5]) was bound to the Protein A Dynabeads. To preclude non-specific binding of the concentrated pN $\alpha 1(V)$ -medium to the Protein A Dynabeads, we performed a titration experiment. Concentrated pN $\alpha 1(V)$ -medium was incubated with Protein A Dynabeads and the optimal protein concentration, was determined by Western blotting after subtraction of non-specific binding. For each co-immunoprecipitation assay, a negative control, which consisted of Protein A Dynabeads incubated with the concentrated pN $\alpha 1(V)$ -medium, was included. Subsequently, concentrated pN $\alpha 1(V)$ -medium was incubated for 4h at 4°C with the Dynabeads-18G5 complex with constant rotation. The Dynabeads-18G5-pN $\alpha 1(V)$ -protein-complex was washed and incubated overnight at 4°C with constant rotation in the presence of cell lysates of HEK293T cells, transfected with either one of the expression plasmids encoding the different potential interacting domains, or in the presence of 10ng of human recombinant TGF- $\beta 1$ or 3 μg affinity purified type I collagen C-propeptide. The resulting complex was washed and eluted with non-reducing lane marker sample buffer (ThermoScientific) supplemented with DTT (Sigma) (except for the type I collagen C-propeptide), boiled and loaded on NuPAGE® 4-12% Bis-Tris gels (Invitrogen) or on a self-poured 10% SDS-PAGE gel for TNC. After SDS-PAGE, Western blotting was performed using the mouse 18G5 antibody [5, 10], the mouse anti-flag M2 antibody (Sigma-Aldrich) for probing the interaction partner proteins and the mouse anti-TGF- $\beta 1$ antibody (TB21, ab27969, Abcam), followed by incubation with the secondary antibody ECLplex goat-anti-mouse IgG-Cy5 (GE Healthcare). For the co-immunoprecipitation assay with the pN $\alpha 1(V)$ -protein and the type I collagen C-propeptide, rabbit antibodies to the CUB1CUB2 fragment of mouse PCPE-1 and guinea-pig antibodies to the mouse type I C-propeptide were used [24]. In order to detect the pN $\alpha 1(V)$ -protein, a rabbit polyclonal antibody directed against type V collagen (ab7046, Abcam) was used. Next, the blot was probed with the secondary antibodies Alexa fluor488 goat anti-rabbit (A-11008, Molecular Probes) and Alexa fluor594 goat-anti-guinea pig (A-11076, Molecular Probes). Scanning of the membranes was performed on a Typhoon 9400 (GE Healthcare). Detection of the type I collagen C-propeptide was performed by chemiluminescence using the SuperSignal West Dura Extended Duration Substrate kit (Thermo Scientific). To investigate the specificity of the co-immunoprecipitation assay, negative control samples, containing proteins which do not interact with the pN $\alpha 1(V)$ -protein (gift of A. Dheedene), were included. These samples consisted of cell lysates of HEK293T cells expressing: SON DNA binding protein; myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, *Drosophila*) translocated to 1 (MLLT1); polycomb group ring finger 2 (PCGF2) and zinc finger CCCH-type containing 4 (ZC3H4), each of which was generated in the flag-tag vector pMG2-p85. In addition, as a negative control the co-immunoprecipitation assays were carried out with the cell lysate of untransfected HEK293T cells (Supplementary Figure 2).

Solid Phase Binding Assay

1 μg /well of purified pepsinized fetal calf type I collagen, and acid soluble and pepsinized type V collagen were coated onto microtiter plates (96 wells, Greiner) overnight at 4°C. The wells were blocked with 3% fat-free milk in PBS for 2 h. Serial dilutions (0-40 $\mu\text{g}/\text{ml}$) of pepsinized human type VI collagen, diluted in PBS containing 0.05% Tween 20 and 3% fat-free milk, were added

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

to the wells and incubated overnight at 4°C. After several washes, bound type VI collagen was revealed with a monoclonal antibody against type VI collagen (5C6, Hybridoma Bank, University of Iowa, USA) followed by incubation with a secondary mouse antibody conjugated to peroxidase (Dako A/S). The signal was detected using 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) as the chromogenic substrate. Absorbance was measured at 405nm with an ELISA microplate reader.

Enzyme digests

2 μ g of the purified pN α 1(V)-protein was incubated overnight at 37°C, without or with 100ng of BMP-1 (1:20 enzyme/substrate ratio) in assay buffer (50mM Tris-HCl pH7.4, 150mM NaCl, 5mM CaCl₂). When indicated, 1.56 μ g of PCPE-1 was added. Digestion products were precipitated with 10% trichloroacetic acid after adding 1% Triton X100. After centrifugation, pellets were washed twice with cold ethanol and acetone, resuspended in Laemmli buffer and analyzed by 7.5% SDS-PAGE under reducing conditions. Proteins were electrotransferred onto polyvinylidene difluoride membranes (Immobilon-P, Millipore) overnight in 10mM CAPS, pH11, 5% methanol. Then, membranes were saturated for 2 hours with 10% non-fat milk in PBS (Sigma-Aldrich). Double immunolabeling using two different colorimetric detection kits was performed to discriminate BMP-1 from the pN α 1(V)-protein and digestion product bands. The membrane was first probed with a primary polyclonal antibody against the TSPN-1 domain [5] followed by incubation with an alkaline phosphatase conjugated goat anti-rabbit secondary antibody (BioRad). Detection of the protein bands was performed using an alkaline phosphatase colorimetric kit (AP Color kit, BioRad). Subsequently, the membrane was incubated with a primary polyclonal antibody against BMP-1 followed by incubation with a peroxidase conjugated goat anti-rabbit secondary antibody (DAKO A/S) and the signal was detected using a peroxidase colorimetric kit (Sigma-Aldrich).

The time course of the processing of the type V procollagen homotrimer by BMP-1 was followed as previously described [25]. Briefly, recombinant type V procollagen (3 μ g) was incubated with 100ng of BMP-1 at 37°C in 48 μ l of 50mM Tris/HCl pH7.5, 150mM NaCl, 5mM CaCl₂, 0.1% Brij-35, in the presence or absence of 200ng of PCPE-1. Aliquots (6.4 μ l) were removed at 2, 4 and 6 h and the reaction was stopped by heating (60°C, 15 min) in Laemmli's loading buffer containing 4% 2-mercaptoethanol. Digestion products were analyzed by 10% SDS-PAGE and silver staining.

Relative band intensities were determined after scanning the original membranes (or the dried gel) on optimized, non-edited images, using TINA image analysis software (ver.2.07d, Raytest isotopenmessgerete, GmbH), after background subtraction.

Results and Discussion

Since the yeast two-hybrid (Y2H) system has been successfully used to identify interactions with extracellular proteins (e.g. type IV collagen, thrombospondin 1, laminin β 3, the NC1 domain of type VII collagen, emilin and microfibril-associated glycoprotein-2 [26]) and to study protein assembly of type X collagen [27], we reasoned that the Y2H assay could be successful in identifying binding partners to the type V collagen N-propeptide. We thus designed a Y2H bait construct pGBKT7-pN α 1(V), which encompassed the entire α 1(V) N-propeptide domain and 11 triplets of the COL1 domain [5]. Taking into account that skin is the major affected tissue in classic EDS patients, we generated a human dermal fibroblast cDNA library.

Screening of this library with the pGBKT7-pN α 1(V) bait vector resulted in the growth of numerous yeast colonies (~ 4000 clones), even after stringent nutritional selection. Due to this high number of positive yeast colonies, a PCR strategy was developed to screen all different inserts. Approximately 32% of the positive yeast colonies (1128 colonies) yielded one specific PCR band, indicating that only one insert, encoding one interaction partner, was present. The

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

remaining positive yeast colonies yielded multiple PCR bands, suggesting that multiple inserts were present; the absence of a PCR band indicates that the PCR conditions were not optimal. Direct sequencing of the 1128 specific PCR products yielded interpretable results for ~60% of the PCR fragments. Sequences that encoded nuclear and cytoskeleton proteins and proteins known as being false positive Y2H interactions (e.g. ribosomal proteins and proteasome subunits, Supplementary Table 2) were omitted from further investigation. This resulted in the retention of nucleotide sequences encoding domains of 12 polypeptide chains/proteins putatively interacting with the $\alpha 1(V)$ -N-propeptide (Table 1). Each interacting domain was identified at least twice. These binding partners included ECM proteins that are involved in matrix assembly and cell-matrix interactions (the different α -chains of types I and VI collagen, fibronectin and tenascin C), and ECM proteins that are involved in collagen biosynthesis and remodeling (procollagen C-proteinase enhancer-1, matrix metalloproteinase-2, tissue inhibitor of matrix metalloproteinases-1, transforming growth factor- $\beta 1$ and N-acetylgalactosaminyltransferase 7). Moreover, through the Y2H assay, information about the interacting domains could be provided (Table 1). However, it should be noted that interactions between the $\alpha 1(V)$ -N-propeptide and these identified proteins might involve multiple regions and also regions that are not included in this Y2H screen since our fibroblast cDNA library does not contain all protein-coding full length sequences. In addition, some interactions might depend on post-translational modifications and/or the structure of mature, processed protein.

To validate the identified interactions, several protein-protein binding assays were applied. The binding of the $\alpha 1(V)$ -N-propeptide to PCPE-1, TGF- $\beta 1$, and types I and VI collagen was confirmed by surface plasmon resonance (SPR) (Figure 2). As was previously shown for the binding of the procollagen C-propeptide to PCPE-1 [22], the binding of the $\alpha 1(V)$ -N-propeptide to PCPE-1 also increased in the presence of 5mM CaCl_2 (Figure 2A). The complexes formed between PCPE-1 and the $\alpha 1(V)$ -N-propeptide were stable as demonstrated by their very low spontaneous dissociation rate (Figure 2A). Soluble TGF- $\beta 1$ bound to the $\alpha 1(V)$ -N-propeptide, but dissociated rapidly (Figure 2B). In addition, SPR arrays showed the formation of very stable complexes of the $\alpha 1(V)$ -N-propeptide with pepsinized type I and VI collagens (Figure 2C). Additionally, we confirmed by co-immunoprecipitation that the $\alpha 1(V)$ -N-propeptide binds to the identified domains of some of the binding partners (FN, $\alpha 1(VI)$ -collagen chain, TIMP-1, MMP-2 and TNC) (Figure 3A), to the full-length TGF- $\beta 1$ and to the C-propeptide of type I collagen (Figure 3B). No interaction was detected with four negative controls (SON, MMLT1, PCGF2 and ZC3H4, not shown). Hence, from the twelve identified proteins that interact with the $\alpha 1(V)$ -N-propeptide, eleven were confirmed by SPR and/or co-immunoprecipitation. Unfortunately, the interaction of the $\alpha 1(V)$ -N-propeptide with the N-acetylgalactosaminyl-transferase 7 (GalNAc-T7) enzyme could not be confirmed since the protein is not commercially available for SPR studies and, for some reason, our different attempts to produce the recombinant protein failed. This enzyme belongs to the family of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (EC 2.4.1.41) [28]. Since this family of enzymes is poorly characterized, the putative physiological significance of this interaction remains to be determined.

Some proteins identified from our Y2H screen were previously reported as type V collagen ligands. This is the case for the $\alpha 1$ - and $\alpha 2$ -chains of type I collagen [1, 29] and FN [30]. However, our findings add information to the previously reported data. First, we show that the binding site for FN in type V collagen was located within the $\alpha 1(V)$ -N-propeptide domain. Second, our Y2H results suggest that the $\alpha 1(V)$ -N-propeptide interacts with the C-propeptide domains of the $\alpha 1(I)$ - and $\alpha 2(I)$ -collagen chains. These interactions may represent an important factor in the growth control of the heterotypic type I/V. The binding of $\alpha 1(V)$ -N-propeptide with the type I collagen C-propeptide domains can directly regulate the rate of the type I procollagen C-propeptide and thus the accretion of heterotypic fibrils by limiting the collagen I molecule deposit. Moreover, although the triple helical domains of the $\alpha 1(I)$ - and $\alpha 2(I)$ -collagen chains

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

were not identified in our Y2H screen, we showed using SPR assays and solid-phase binding assays that the type I collagen triple helix domain may also be involved in the type I – type V collagen interaction (Figure 2C and 3A). The fact that this interacting protein domain is absent from our Y2H screen suggests that the triple helix conformation is likely a prerequisite for this interaction to occur. Another possibility is that the sequence encoding this domain was not withheld due to our selection procedure.

New binding partners were also identified. We demonstrated for the first time a direct interaction between the $\alpha 1(V)$ -N-propeptide and type VI collagen (Figure 2C). The most common type VI collagen form is the heterotrimer $\alpha 1(VI)\alpha 2(VI)\alpha 3(VI)$ (encoded by *COL6A1*, *COL6A2* and *COL6A3* genes respectively) [31]. Each α -chain consists of a central short collagenous domain flanked by several von Willebrand factor type A (vWA) modules at the N- and C-terminus [31]. Our Y2H data revealed that different regions of the type VI collagen molecule are involved in the $\alpha 1(V)$ -N-propeptide interaction. Whereas the $\alpha 1(VI)$ - and $\alpha 3(VI)$ -chains interact with the $\alpha 1(V)$ -N-propeptide through their globular domains (Table 1), the $\alpha 2(VI)$ -chain interacts with the $\alpha 1(V)$ -N-propeptide through its triple helix (Table 1). We showed that the $\alpha 1(V)$ -N-propeptide binding sites are retained in the pepsinized form of type VI collagen (Figure 2C). Treatment with pepsin usually removes the globular domains of collagens. However, it has been shown that pepsinized type VI collagen consists of a mixture of dimers and tetramers which retain small globular domains at the extremities of the molecules [32], suggesting that the binding sites located within the globular domains are likely present in the pepsinized molecule. Moreover, using a solid phase binding assay, we investigated if pepsinized type VI collagen can bind to the intact form of type V collagen extracted from tissues with acetic acid (Figure 4). Mature type V collagen retains the VAR, COL2 and NC2 domains, while the TSPN-1 domain is removed by BMP-1 during processing (Figure 1). When pepsinized type VI collagen was used as soluble ligand, a strong and saturable binding was obtained with acid-soluble type V collagen (Figure 4A). A weaker binding of pepsinized type VI collagen resulted with type I collagen. In contrast, no binding to pepsinized type V collagen was detected (Figure 4A). These results underscore the Y2H data and show that type VI collagen binds to mature type V collagen found in tissues (Figure 4C).

Interestingly, although we used the monomeric $\alpha 1(V)$ -N-propeptide in our initial Y2H screen, we show by our solid phase binding assay that the heterotrimeric type V collagen interacts with type VI collagen, thereby supporting the interaction of type VI collagen with the physiological form of type V collagen. Previously, type VI collagen was shown to interact with various ECM components including type I, II, IV, and XIV collagens, FN, fibromodulin, hyaluronic acid, perlecan, decorin, biglycan, tenascin-X and cell receptors such as the integrins and the NG2 transmembrane proteoglycan [33]. Type VI collagen is widely distributed in connective tissues and particularly abundant in the pericellular environment, where it functions to anchor the basement membrane to the underlying connective tissue [31]. Along this line, ultrastructural studies suggested that types V and VI collagen form intricate networks in the dermis that can be linked to collagen fibrils, elastic fibers and the basal lamina [34]. As such, the type V - type VI collagen protein complex may serve as a molecular bridge in the cell-matrix environment, and may be essential to maintain the architecture of the dermal matrix. Interestingly, Ullrich congenital muscular dystrophy (UCMD), which is caused by mutations in *COL6A1*, *COL6A2* or *COL6A3*, shows clinical and ultrastructural characteristics overlapping with classic EDS, such as alterations of collagen fibril morphology (variation in size, loose packaging of fibrils and composite fibers) and increase in ground substance [35]. In view of the identified type V – type VI collagen interactions, cutaneous manifestations in UCMD may be mediated by disturbance of type V collagen function. In addition, as defects in type VI collagen are found only in approximately 40% of the UCMD patients and since the involvement of type VI collagen has been excluded in a number of UCMD cases [35], type V collagen might be a candidate for

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

unexplained cases of UCMD. Vice versa, type VI collagen may be a good candidate gene for unresolved (classic) EDS patients.

Our findings also show that the $\alpha 1(V)$ -N-propeptide interacts with PCPE-1, a potent protein enhancer of the C-terminal processing of type I, II and III procollagens by members of procollagen C-proteinase family, notably the bone morphogenetic protein-1 (BMP-1) [36]. However, BMP-1 cleaves the type V procollagen homotrimer at both ends, releasing very rapidly the TSPN-1 domain of the N-propeptide and more slowly the C-propeptide [25]. To examine the function of this interaction, PCPE-1 enhancement of the N-propeptide cleavage by BMP-1 was assessed using an *in vitro* enzymatic assay. Recombinant $\alpha 1(V)$ -N-propeptide was incubated with BMP-1 in the presence or absence of PCPE-1 (Figure 5A) and the reaction products were detected by immunoblotting. Densitometric measurements of the intensities of the bands corresponding to the TSPN-1 domain showed that the calculated ratio of the TSPN-1 band intensity in the presence of BMP-1 and PCPE-1 (Figure 5A, lane 4) over the TSPN-1 band intensity with BMP-1 alone (Figure 5A, lane 2) was 1.9. Subsequently, we performed a time course study using recombinant type V procollagen homotrimer as a substrate. The rates of release of the $\alpha 1(V)$ -N-propeptide TSPN-1 fragment by BMP-1 in the presence or absence of PCPE-1 were compared after 2, 4 and 6 hours of incubation (Figure 5B). The intensity of the TSPN-1 band in the presence of PCPE-1 increased progressively with time (Figure 5B) and was higher than in its absence. Densitometric measurements for the BMP-1 digests indicated a 1.7, 2.5 and 2.4 fold enhancement in the presence of PCPE-1 at 2, 4 and 6 hours of incubation, respectively. These results are consistent with the ratio obtained in Figure 5A. Our study clearly indicates that PCPE-1 binding to the $\alpha 1(V)$ -N-propeptide is required for the enhancement of BMP-1 activity, hereby providing a new biological function for PCPE-1. This finding was unexpected because thus far PCPE-1 enhancing activity was thought to be restricted to the procollagen C-proteinase activity of BMP-1 [37]. Moreover, using a different approach, it has been previously reported that PCPE-1 has no effect on the *in vitro* BMP-1 processing of type V procollagen N-propeptide [37]. Consistent with our *in vitro* enzymatic activity, the identified binding sites are located within the PCPE CUB domains, also involved in the PCPE binding of fibrillar collagen C-propeptides [37]. As the type V collagen N-propeptide domain is of particular importance in regulating the collagen fibril diameter, we hypothesize that PCPE-1 can directly modulate collagen fibrillogenesis by regulating type V collagen N-propeptide processing.

Additional newly identified binding partners of the $\alpha 1(V)$ -N-propeptide were TGF- β 1 and MMP-2, the most widespread matrix metalloproteinase, and its potent tissue inhibitor TIMP-1. TGF- β 1 is one of the most pleiotropic and multifunctional growth factors which upregulates the expression of several ECM components (FN, biglycan, types I and V collagen, TNC) [38, 39]. Additionally, it was shown that *COL5A1* is a target of TGF- β 1 signaling with a potential role in osteogenesis [39]. Activation of the TGF- β 1 molecule results in the stimulation of many of its own activators, including furin [40], which is involved in the processing of the C-propeptide of the pro $\alpha 1(V)$ -collagen chain [4]. Here we showed that, besides these regulatory actions of TGF- β 1 on *COL5A1*, also a physical interaction between the two proteins exists. The downstream effects of this interaction so far remain elusive. However, it emphasizes further the complexity of the relation between TGF- β 1 signaling and the ECM. Moreover, it has previously been shown that TGF- β 1 binds to the cysteine-rich repeat (CRR) domain of the type II collagen N-propeptide [41]. Proteolytic removal of this CRR domain releases TGF- β 1 in cartilage and regulates chondrogenesis. Taken together, our findings suggest a novel function for the $\alpha 1(V)$ -N-propeptide domain. As type V procollagen is synthesized, processed and deposited in the ECM, the $\alpha 1(V)$ -N-propeptide could serve as a reservoir controlling the availability of TGF- β 1. The possibility that TGF- β 1 can bind to the $\alpha 1(V)$ -TSPN-1 domain, whose release is mediated by BMP-1, suggests that this interaction may control ECM assembly and homeostasis and needs further investigation. Remodelling of the ECM depends on the finely-tuned regulation of the MMPs and their endogenous inhibitors TIMPs, both identified as binding partners of $\alpha 1(V)$ -N-

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

propeptide. Various MMP-2 substrates have been identified including the denatured collagen proteins [42], the $\alpha 3(VI)$ -collagen chain [43] and the latency associated peptide of TGF- $\beta 1$ [40] and the native type V collagen [44]. Previously, it was shown that the level of MMP-2 is up- or down-regulated depending on the amount of type V collagen in contact with fibroblasts [44]. Here, we showed that MMP-2 binds to the $\alpha 1(V)$ -N-propeptide through the MMP-2 hemopexin-like domain which is involved in substrate specificity (Table 1). A recent large-scale Y2H screening revealed that MMP-2 interacts with TGF- $\beta 1$ and with the pro $\alpha 1(I)$ -collagen chain [45]. Furthermore, TGF- $\beta 1$ was involved in the induction of TIMP-1 expression [46]. Additionally, it was demonstrated that, at high concentration, the $\alpha 2(VI)$ -collagen chain formed very stable complexes with proMMP-2 [47]. As such, the type V/VI collagen complex may also be important in modulating MMP-2 availability in the ECM.

In conclusion, we present novel interactions between the $\alpha 1(V)$ -N-propeptide and proteins that thus far, have not been implicated in collagen fibrillogenesis and/or additional functions of type V collagen. These results may shed new light on the formation of heterotypic type I/V collagen fibrils and on the pathogenetic mechanism through which type V collagen defects lead to classic EDS. Furthermore, the discovery of new binding partners of type V collagen now allows the identification of new candidate genes for this connective tissue disorder. In agreement with this, mutations in type I collagen [48, 49] and fibronectin [50], two ligands identified in our Y2H screen, have been reported in some cases of EDS (EDS linked to *COL1A1*: MIM#130000, MIM#130060, EDS linked to *COL1A2*: MIM#225320, MIM#130060, EDS linked to *FN1*: MIM#225310). The genes encoding the identified proteins can now be screened as potential candidates in patients with (classic) EDS or related conditions in whom the underlying genetic defect remains elusive. We characterized in more detail the interaction of the $\alpha 1(V)$ -N-propeptide with type VI collagen and PCPE-1 respectively. We postulate that the type V – type VI collagen interaction is of particular importance for maintaining the ECM architecture and, most importantly, that disturbance of this interaction might explain the overlap in ultrastructural abnormalities observed between UCMD and EDS. In addition, we showed that PCPE-1 is capable of enhancing the rate of cleavage of the $\alpha 1(V)$ -N-propeptide by BMP-1. As such, it is possible that defects in PCPE-1 disturb normal processing of fibrillar collagen proteins and subsequent collagen fibrillogenesis and contribute in this way to the pathogenesis of (classic) EDS.

Acknowledgements

We thank Dr. D. Hulmes (UMR 5086, CNRS - University Lyon 1) and Dr. P. Findell (Roche-BioScience) for kindly providing recombinant PCPE-1 protein, Dr. T. Sasaki (MPI, Martinsried, Germany) for kindly providing human pepsinized type VI collagen, and A. Dheedene for kindly providing SON, MLLT1, PCGF2 and ZC3H4. We thank E. De Vogelaere, J. Weytens and P. Tapaneeyaphan for excellent technical assistance with co-immunoprecipitation assays. We are also grateful to C. Faye for excellent technical assistance with SPR arrays and to C. Marquette and L. Blum (ICBMS, UMR 5246 CNRS, University Lyon 1, France) for providing access to the spotter. We thank Prof. J. Tavernier and Dr. S. Lievens (VIB Department of Medical Protein Research and Ghent University Department of Biochemistry, Belgium) for providing the pMET7-fC-SH2 and pMG2-p85 vectors. We are grateful to the IFR128 Biosciences-Gerland core facilities for the use of the Biacore T100 and the Flexchip apparatus ("Production et Analyse des Protéines") and the Nanodrop ("Analyse génétique") and for their helpful technical assistance. The support of the GE Healthcare staff (C. Quetard, P. Linden, T. Salomon, X. Rousselot and A. Sylvan) in the Biacore Flexchip experiments is thankfully acknowledged. We are grateful to Dr. S. Kotev-Emeth (Goldschleger Eye Institute, Tel Aviv University faculty of Medicine, Sheba Medical Center) for help with the densitometry experiments.

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins**Funding**

This work was supported by grant G.0171.05 from the Fund for Scientific Research-Flanders to A. De Paepe, a Methusalem grant 08/01M01108 from the Ghent University to A. De Paepe and by the "Geneskin" consortium within the sixth framework program of the European Commission (number 512117), and by a grant GIS Maladies Rares (INSERM) and the Emergence Program of the Région Rhône-Alpes to F. Ruggiero and S. Ricard-Blum, and by grants 736/01 and 1360/07 from the Israel Science Foundation to E. Kessler. F.M. is a postdoctoral fellow of the Fund for Scientific Research-Flanders.

References

- 1 Birk, D. E. (2001) Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron*. **32**, 223-237
- 2 Chanut-Delalande, H., Bonod-Bidaud, C., Cogne, S., Malbouyres, M., Ramirez, F., Fichard, A. and Ruggiero, F. (2004) Development of a functional skin matrix requires deposition of collagen V heterotrimers. *Mol. Cell Biol.* **24**, 6049-6057
- 3 Wenstrup, R. J., Florer, J. B., Davidson, J. M., Phillips, C. L., Pfeiffer, B. J., Menezes, D. W., Chervoneva, I. and Birk, D. E. (2006) Murine model of the Ehlers-Danlos syndrome. *col5a1* haploinsufficiency disrupts collagen fibril assembly at multiple stages. *J. Biol. Chem.* **281**, 12888-12895
- 4 Unsold, C., Pappano, W. N., Imamura, Y., Steiglitz, B. M. and Greenspan, D. S. (2002) Biosynthetic processing of the pro- $\alpha 1(V)2$ pro- $\alpha 2(V)$ collagen heterotrimer by bone morphogenetic protein-1 and furin-like proprotein convertases. *J. Biol. Chem.* **277**, 5596-5602
- 5 Bonod-Bidaud, C., Beraud, M., Vaganay, E., Delacoux, F., Font, B., Hulmes, D. J. and Ruggiero, F. (2007) Enzymatic cleavage specificity of the pro $\alpha 1(V)$ chain processing analysed by site-directed mutagenesis. *Biochem. J.* **405**, 299-306
- 6 Colige, A., Ruggiero, F., Vandenberghe, I., Dubail, J., Kesteloot, F., Van Beeumen, J., Beschin, A., Brys, L., Lapiere, C. M. and Nusgens, B. (2005) Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I-III and V. *J. Biol. Chem.* **280**, 34397-34408
- 7 Hausser, I. and Anton-Lamprecht, I. (1994) Differential ultrastructural aberrations of collagen fibrils in Ehlers-Danlos syndrome types I-IV as a means of diagnostics and classification. *Hum. Genet.* **93**, 394-407
- 8 Malfait, F., Coucke, P., Symoens, S., Loeys, B., Nuytinck, L. and De Paepe, A. (2005) The molecular basis of classic Ehlers-Danlos syndrome: a comprehensive study of biochemical and molecular findings in 48 unrelated patients. *Hum. Mutat.* **25**, 28-37
- 9 Mitchell, A. L., Schwarze, U., Jennings, J. F. and Byers, P. H. (2009) Molecular mechanisms of classical Ehlers-Danlos syndrome (EDS). *Hum. Mutat.* **30**, 995-1002
- 10 Symoens, S., Malfait, F., Renard, M., Andre, J., Hausser, I., Loeys, B., Coucke, P. and De Paepe, A. (2009) COL5A1 signal peptide mutations interfere with protein secretion and cause classic Ehlers-Danlos syndrome. *Hum. Mutat.* **30**, E395-403
- 11 Giunta, C., Nuytinck, L., Raghunath, M., Hausser, I., De Paepe, A. and Steinmann, B. (2002) Homozygous Gly530Ser substitution in COL5A1 causes mild classical Ehlers-Danlos syndrome. *Am. J. Med. Genet.* **109**, 284-290
- 12 Takahara, K., Schwarze, U., Imamura, Y., Hoffman, G. G., Toriello, H., Smith, L. T., Byers, P. H. and Greenspan, D. S. (2002) Order of intron removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in abnormal pro- $\alpha 1(V)$ N-propeptides and Ehlers-Danlos syndrome type I. *Am. J. Hum. Genet.* **71**, 451-465
- 13 Fichard, A., Kleman, J. P. and Ruggiero, F. (1995) Another look at collagen V and XI molecules. *Matrix Biol.* **14**, 515-531

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

- 14 Delacoux, F., Fichard, A., Cogne, S., Garrone, R. and Ruggiero, F. (2000) Unraveling the amino acid sequence crucial for heparin binding to collagen V. *J. Biol. Chem.* **275**, 29377-29382
- 15 Whinna, H. C., Choi, H. U., Rosenberg, L. C. and Church, F. C. (1993) Interaction of heparin cofactor II with biglycan and decorin. *J. Biol. Chem.* **268**, 3920-3924
- 16 Xie, R. L. and Long, G. L. (1996) Elements within the first 17 amino acids of human osteonectin are responsible for binding to type V collagen. *J. Biol. Chem.* **271**, 8121-8125
- 17 Lethias, C., Carisey, A., Comte, J., Cluzel, C. and Exposito, J. Y. (2006) A model of tenascin-X integration within the collagenous network. *FEBS Lett.* **580**, 6281-6285
- 18 Ruggiero, F., Champliand, M. F., Garrone, R. and Aumailley, M. (1994) Interactions between cells and collagen V molecules or single chains involve distinct mechanisms. *Exp. Cell Res.* **210**, 215-223
- 19 Vogel, W., Gish, G. D., Alves, F. and Pawson, T. (1997) The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol. Cell.* **1**, 13-23
- 20 Behrendt, N., Jensen, O. N., Engelholm, L. H., Mortz, E., Mann, M. and Dano, K. (2000) A urokinase receptor-associated protein with specific collagen binding properties. *J. Biol. Chem.* **275**, 1993-2002
- 21 Van Criekege, W. and Beyaert, R. (1999) Yeast Two-Hybrid: State of the Art. *Biol. Proced. Online.* **2**, 1-38
- 22 Ricard-Blum, S., Bernocco, S., Font, B., Moali, C., Eichenberger, D., Farjanel, J., Burchardt, E. R., van der Rest, M., Kessler, E. and Hulmes, D. J. (2002) Interaction properties of the procollagen C-proteinase enhancer protein shed light on the mechanism of stimulation of BMP-1. *J. Biol. Chem.* **277**, 33864-33869
- 23 Faye, C., Chautard, E., Olsen, B. R. and Ricard-Blum, S. (2009) The first draft of the endostatin interaction network. *J. Biol. Chem.* **284**, 22041-22047
- 24 Kessler, E. and Adar, R. (1989) Type I procollagen C-proteinase from mouse fibroblasts. Purification and demonstration of a 55-kDa enhancer glycoprotein. *Eur. J. Biochem.* **186**, 115-121
- 25 Kessler, E., Fichard, A., Chanut-Delalande, H., Brusel, M. and Ruggiero, F. (2001) Bone morphogenetic protein-1 (BMP-1) mediates C-terminal processing of procollagen V homotrimer. *J. Biol. Chem.* **276**, 27051-27057
- 26 Penner, A. S., Rock, M. J., Kielty, C. M. and Shipley, J. M. (2002) Microfibril-associated glycoprotein-2 interacts with fibrillin-1 and fibrillin-2 suggesting a role for MAGP-2 in elastic fiber assembly. *J. Biol. Chem.* **277**, 35044-35049
- 27 Nielsen, V. H., Bendixen, C., Arnbjerg, J., Sorensen, C. M., Jensen, H. E., Shukri, N. M. and Thomsen, B. (2000) Abnormal growth plate function in pigs carrying a dominant mutation in type X collagen. *Mamm. Genome.* **11**, 1087-1092
- 28 Bennett, E. P., Hassan, H., Hollingsworth, M. A. and Clausen, H. (1999) A novel human UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett.* **460**, 226-230
- 29 Chanut-Delalande, H., Fichard, A., Bernocco, S., Garrone, R., Hulmes, D. J. and Ruggiero, F. (2001) Control of heterotypic fibril formation by collagen V is determined by chain stoichiometry. *J. Biol. Chem.* **276**, 24352-24359
- 30 Zoppi, N., Gardella, R., De Paepe, A., Barlati, S. and Colombi, M. (2004) Human fibroblasts with mutations in COL5A1 and COL3A1 genes do not organize collagens and fibronectin in the extracellular matrix, down-regulate $\alpha 2\beta 1$ integrin, and recruit $\alpha v\beta 3$ instead of $\alpha 5\beta 1$ integrin. *J. Biol. Chem.* **279**, 18157-18168
- 31 Colombatti, A., Mucignat, M. T. and Bonaldo, P. (1995) Secretion and matrix assembly of recombinant type VI collagen. *J. Biol. Chem.* **270**, 13105-13111

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

- 32 Tillet, E., Ruggiero, F., Nishiyama, A. and Stallcup, W. B. (1997) The membrane-spanning proteoglycan NG2 binds to collagens V and VI through the central nonglobular domain of its core protein. *J. Biol. Chem.* **272**, 10769-10776
- 33 Higashi, K., Higuchi, I., Niiyama, T., Uchida, Y., Shiraishi, T., Hashiguchi, A., Saito, A., Horikiri, T., Suehara, M., Arimura, K. and Osame, M. (2006) Abnormal expression of proteoglycans in Ullrich's disease with collagen VI deficiency. *Muscle Nerve*. **33**, 120-126
- 34 Kobayasi, T. and Karlsmark, T. (2006) Type V and VI collagen for cohesion of dermal fibrillar structures. *J. Submicrosc. Cytol. Pathol.* **38**, 103-108
- 35 Kirschner, J., Hausser, I., Zou, Y., Schreiber, G., Christen, H. J., Brown, S. C., Anton-Lamprecht, I., Muntoni, F., Hanefeld, F. and Bonnemann, C. G. (2005) Ullrich congenital muscular dystrophy: connective tissue abnormalities in the skin support overlap with Ehlers-Danlos syndromes. *Am. J. Med. Genet. A*. **132**, 296-301
- 36 Kessler, E., Takahara, K., Biniaminov, L., Brusel, M. and Greenspan, D. S. (1996) Bone morphogenetic protein-1: the type I procollagen C-proteinase. *Science*. **271**, 360-362
- 37 Moali, C., Font, B., Ruggiero, F., Eichenberger, D., Rousselle, P., Francois, V., Oldberg, A., Bruckner-Tuderman, L. and Hulmes, D. J. (2005) Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1. *J. Biol. Chem.* **280**, 24188-24194
- 38 Gleizes, P. E., Munger, J. S., Nunes, I., Harpel, J. G., Mazzieri, R., Noguera, I. and Rifkin, D. B. (1997) TGF-beta latency: biological significance and mechanisms of activation. *Stem Cells*. **15**, 190-197
- 39 Kahai, S., Vary, C. P., Gao, Y. and Seth, A. (2004) Collagen, type V, alpha1 (COL5A1) is regulated by TGF-beta in osteoblasts. *Matrix Biol.* **23**, 445-455
- 40 Stawowy, P., Margeta, C., Kallisch, H., Seidah, N. G., Chretien, M., Fleck, E. and Graf, K. (2004) Regulation of matrix metalloproteinase MT1-MMP/MMP-2 in cardiac fibroblasts by TGF-beta1 involves furin-convertase. *Cardiovasc. Res.* **63**, 87-97
- 41 Zhu, Y., Oganessian, A., Keene, D. R. and Sandell, L. J. (1999) Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the extracellular matrix of prechondrogenic tissue and binds to TGF-beta1 and BMP-2. *J. Cell Biol.* **144**, 1069-1080
- 42 Gunther, U., Schuppan, D., Bauer, M., Matthes, H., Stallmach, A., Schmitt-Graff, A., Riecken, E. O. and Herbst, H. (1999) Fibrogenesis and fibrolysis in collagenous colitis. Patterns of procollagen types I and IV, matrix-metalloproteinase-1 and -13, and TIMP-1 gene expression. *Am. J. Pathol.* **155**, 493-503
- 43 Myint, E., Brown, D. J., Ljubimov, A. V., Kyaw, M. and Kenney, M. C. (1996) Cleavage of human corneal type VI collagen alpha 3 chain by matrix metalloproteinase-2. *Cornea*. **15**, 490-496
- 44 Kerkvliet, E. H., Jansen, I. C., Schoenmaker, T., Beertsen, W. and Everts, V. (2003) Collagen type I, III and V differently modulate synthesis and activation of matrix metalloproteinases by cultured rabbit periosteal fibroblasts. *Matrix Biol.* **22**, 217-227
- 45 Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamasas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P. and Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature*. **437**, 1173-1178
- 46 Gomez, D. E., Alonso, D. F., Yoshiji, H. and Thorgeirsson, U. P. (1997) Tissue inhibitors of metalloproteinases: structure, regulation and biological functions. *Eur. J. Cell. Biol.* **74**, 111-122

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

- 47 Freise, C., Erben, U., Muche, M., Farndale, R., Zeitz, M., Somasundaram, R. and Ruehl, M. (2009) The alpha 2 chain of collagen type VI sequesters latent proforms of matrix-metalloproteinases and modulates their activation and activity. *Matrix Biol.* **28**, 480-489
- 48 Byers, P. H., Duvic, M., Atkinson, M., Robinow, M., Smith, L. T., Krane, S. M., Grealley, M. T., Ludman, M., Matalon, R., Pauker, S., Quanbeck, D. and Schwarze, U. (1997) Ehlers-Danlos syndrome type VIIA and VIIB result from splice-junction mutations or genomic deletions that involve exon 6 in the COL1A1 and COL1A2 genes of type I collagen. *Am. J. Med. Genet.* **72**, 94-105
- 49 Malfait, F., Symoens, S., De Backer, J., Hermanns-Le, T., Sakalihasan, N., Lapiere, C. M., Coucke, P. and De Paepe, A. (2007) Three arginine to cysteine substitutions in the pro-alpha (I)-collagen chain cause Ehlers-Danlos syndrome with a propensity to arterial rupture in early adulthood. *Hum. Mutat.* **28**, 387-395
- 50 Arneson, M. A., Hammerschmidt, D. E., Furcht, L. T. and King, R. A. (1980) A new form of Ehlers-Danlos syndrome. Fibronectin corrects defective platelet function. *JAMA.* **244**, 144-147
- 51 Lampe, A. K. and Bushby, K. M. (2005) Collagen VI related muscle disorders. *J. Med. Genet.* **42**, 673-685

Figure legends

Figure 1. Type V collagen. A. Schematic representation of the pro $\alpha 1(V)$ - and pro $\alpha 2(V)$ -collagen chain. The pro $\alpha 1(V)$ -N-propeptide consists of several domains: the signal peptide (SP), the TSPN-1 (or PARP) domain, followed by the VAR, COL2 and NC2 domain. The TSPN-1 and VAR domain together form the NC3 domain. The pN $\alpha 1(V)$ -construct (underlined) encompasses the entire $\alpha 1(V)$ -N-propeptide and 11 triplets of the COL1 triple helix domain. Processing of the $\alpha 1(V)$ -collagen chain is complex and can involve different enzymes. Positions of the BMP-1, ADAMTS-2 and furin cleavage sites are shown (arrows). The location of the previously identified mutations, p.G530S and IVS4-2A>G, are indicated by a black rectangle and a black oval respectively. The pro $\alpha 2(V)$ -collagen chain consists of SP domain, followed by the CR (cysteine rich) domain and a small COL2 domain. The $\alpha 2(V)$ -N-propeptide is apparently not processed by any protease. B. Schematic overview of the mature [$\alpha 1(V)$] $_2\alpha 2(V)$ heterotrimer. The pro $\alpha 1(V)$ -N-propeptide and pro $\alpha 2(V)$ -C-propeptide are processed by BMP-1, and the pro $\alpha 1(V)$ -C-propeptide is processed by furin and/or BMP-1.

Figure 2. Surface plasmon resonance binding assays. A. PCPE-1 was immobilized onto a CM5 sensor chip. Soluble pN $\alpha 1(V)$ -protein was injected in the presence or absence of 5 mM CaCl₂ (dashed line: with CaCl₂, solid line: without CaCl₂). In both instances strong interaction complexes were formed, but the binding of the pN $\alpha 1(V)$ -protein to PCPE-1 increased even more in the presence of CaCl₂. B. Soluble TGF- β 1 was injected over the pN $\alpha 1(V)$ -protein, which was immobilized onto a CM5 sensor chip. Although an interaction occurred, this complex dissociated rapidly. C. The pN $\alpha 1(V)$ -protein was spotted onto a Gold affinity chip and type I (solid line) and type VI (dashed line) collagens were injected. For both collagens a stable interaction with the pN $\alpha 1(V)$ -protein was detected.

Figure 3. Co-immunoprecipitation assay. A. The 18G5 antibody was bound on Protein A Dynabeads. The pN $\alpha 1(V)$ -protein was immobilized on the Protein A Dynabeads-18G5 antibody complex. The Protein A Dynabeads-18G5- pN $\alpha 1(V)$ -protein complex was incubated with cell lysates of HEK293T cells, which expressed flag-tagged domains of different identified interacting proteins ($\alpha 1(VI)$ -collagen chain, FN, TIMP-1, MMP-2 and TNC). The interaction complexes were eluted and evaluated by SDS-PAGE and Western Blotting. The blot was

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

developed using the 18G5 antibody (probed to the pN $\alpha 1(V)$ -protein) and the anti-flag M2 antibody (probed against the flag-tag epitope). For these specific interacting domains an interaction with the pN $\alpha 1(V)$ -protein was shown. In the control sample, no non-specific interaction of the pN $\alpha 1(V)$ -protein with the Protein A Dynabeads is detected. B. The pN $\alpha 1(V)$ -protein was immobilized on Protein A Dynabeads by the 18G5 antibody. The Protein A Dynabeads-18G5-pN $\alpha 1(V)$ -protein complex was incubated with either recombinant human TGF- $\beta 1$ or affinity-purified type I collagen C-propeptide. These interaction complexes were eluted and evaluated by SDS-PAGE and Western Blotting. The pN $\alpha 1(V)$ -protein was detected by a rabbit antibody to type V collagen, while a mouse antibody to TGF- $\beta 1$ and a guinea-pig antibody to mouse type I procollagen C-propeptide were used to reveal TGF- $\beta 1$ and the type I collagen C-propeptide. Both proteins (domains) could be revealed, thereby confirming the interaction between the pN $\alpha 1(V)$ -protein and TGF- $\beta 1$ and the type I collagen C-propeptide respectively. We also used a rabbit antibody to the CUB1CUB2 domains of PCPE-1, which revealed a band corresponding to PCPE-1. An explanation for this can be that this band originates from the culture medium of the HEK293T-EBNA cells used to express the pN $\alpha 1(V)$ -protein, known to express PCPE-1 (Supplementary Figure 2).

Fraction 1 = supernatant of the Protein A Dynabeads after incubation with pN $\alpha 1(V)$ -protein containing culture medium of HEK293T-cells expressing the pN $\alpha 1(V)$ -protein. Elution = elute containing the 18G5 antibody, the pN $\alpha 1(V)$ -protein and the putative interaction partner. * = protein band corresponding to the heavy chain of the 18G5 antibody, ** = protein band corresponding to the light chain of the 18G5 antibody.

Figure 4. A. Solid phase assay. Purified fetal calf type I collagen (diamond), as well as acid-soluble (triangle) and pepsinized (square) type V collagen were immobilized onto a 96-well plate and incubated with serial dilutions of soluble pepsinized type VI collagen. Binding was detected using a monoclonal antibody against human type VI collagen (5C6), followed by incubation with a peroxidase labelled secondary mouse antibody. While a strong interaction of pepsinized type VI collagen was seen with acid-soluble type V collagen, its interaction with pepsinized type V collagen was negligible. This implies that the binding of type VI collagen to type V collagen may be confined to the type V collagen N-propeptide. A weaker interaction was seen with fetal calf type I collagen. B. Schematic structure of the type VI collagen tetramer (adapted from [51]). The regions of the $\alpha 1(VI)$ -, $\alpha 2(VI)$ - and $\alpha 3(VI)$ -collagen chains, as deduced from our Y2H screen, are indicated (arrows). C. Proposed working model for the interaction between type V - type VI collagen. While the type V collagen triple helix is buried within the type I collagen fibrils, its N-propeptide projects through the gap zone of the type I collagen fibril thereby placing its globular domains onto the collagen fibril surface. These globular domains can interact with different regions of type VI collagen molecule.

Figure 5. Effect of PCPE-1 on BMP-1 cleavage of recombinant pN $\alpha 1(V)$ -protein (A) and homotrimeric type V procollagen (B). A. Recombinant pN $\alpha 1(V)$ -protein was incubated with BMP-1 in the absence (lane 2) or presence (lane 4) of PCPE-1. The cleavage product of the pN $\alpha 1(V)$ -protein by BMP-1 that corresponds to the TSPN-1 N-terminal domain of the pro $\alpha 1(V)$ chain is detected by Western Blot analysis. Lanes 1 and 3 correspond to negative controls, incubations were performed without BMP-1 and PCPE-1 (lane 1) and without BMP-1 (lane 3) respectively. The upper unidentified band observed in all samples, whose migration, presence and intensity remains unchanged in all of the lanes, attests for the equal loading of proteins. B. Time course of recombinant type V procollagen homotrimer processing by BMP-1. The rates of release of the TSPN-1 domain from the $\alpha 1(V)$ -N-propeptide by BMP-1 at the indicated time points (0, 2, 4 and 6 hours), in the presence (+) or absence (-) of PCPE-1 is compared. Silver stained SDS-PAGE analysis showed that the band corresponding to the TSPN-1 domain is more intense in the presence of PCPE-1 than in its absence. These results show that the activity of BMP-1 is

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

enhanced by PCPE-1, regardless of the substrate used, respectively the $\alpha 1(V)$ -N-propeptide monomer (A) or the type V procollagen homotrimer (B).

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101061

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins**Table 1. Proteins putatively interacting with the $\alpha 1(V)$ -N-propeptide.**

Protein	# clones	Gene symbol	GenBank Accession number	Position	Interacting domain*
$\alpha 1(I)$ -collagen chain	15	<i>COL1A1</i>	NC_000017.9	AZ 1296 – 1464	C-terminal propeptide
$\alpha 2(I)$ -collagen chain	7	<i>COL1A2</i>	NC_000007.12	AZ 1228 – 1366	C-terminal propeptide
Fibronectin (FN)	2	<i>FN1</i>	NC_000002.10	AZ 1100 - 1292	FN type-III 6 - FN type-III 7 - FN type-III 8
$\alpha 1(VI)$ -collagen chain	2	<i>COL6A1</i>	NC_000021.7	AZ 660 – 845	C-terminal globular domain with: vWF type A repeat 2 (C1 domain) & vWF type A repeat 3 (C2 domain)
$\alpha 2(VI)$ -collagen chain	2	<i>COL6A2</i>	NC_000021.7	AZ 399 – 535	Triple helix domain with 4 potential cell attachment site motifs (RGD2)
$\alpha 3(VI)$ -collagen chain	2	<i>COL6A3</i>	NC_000002.10	AZ1426 – 1865	N-terminal globular domain with: vWF type A repeat 8 (N3 domain) & vWF type A repeat 9 (N2 domain)
Tenascin C (TNC)	2	<i>TNC</i>	NC_000009.10	AZ 1483 – 1622	FN type-III 10 - FN type-III 11
Procollagen C-proteinase enhancer 1 (PCPE-1)	2	<i>PCPE1</i>	NC_000007.12	AZ 39 – 181	CUB1 & CUB2
Tissue inhibitor of metalloproteinases-1 (TIMP-1)	3	<i>TIMP1</i>	NC_000023.9	AZ 103 – 207	NTR
Matrix metalloproteinase-2 (MMP-2)	5	<i>MMP2</i>	NC_000016.8	AZ 346 – 660	FN type II 3 - Hemopexin-like
Transforming growth factor $\beta 1$ (TGF- $\beta 1$)	2	<i>TGFβ1</i>	NC_000019.8	AZ 1 – 390	Full length
N-acetylgalactosaminyl-transferase-7	2	<i>GalNAc-T7</i>	NC_000004.10	AZ 1 – 657	Full length

* As deduced by direct sequencing of the inserts of the Y2H pGADT7-Rec prey vector.

Abbreviations: FN type-III, fibronectin type-III repeat; vWF type A repeat, von Willebrandt factor type A repeat; CUB, C1r/C1s, Uegf and BMP-1 domain; NTR, netrin-1 domain

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

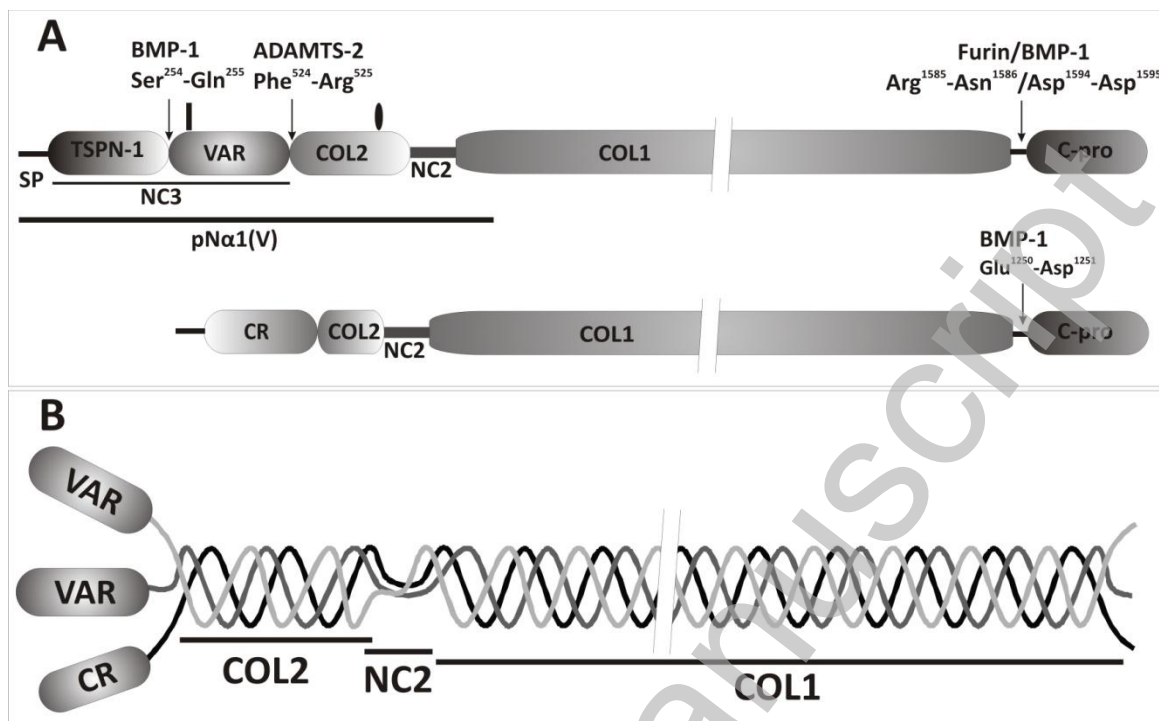


Figure 1

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

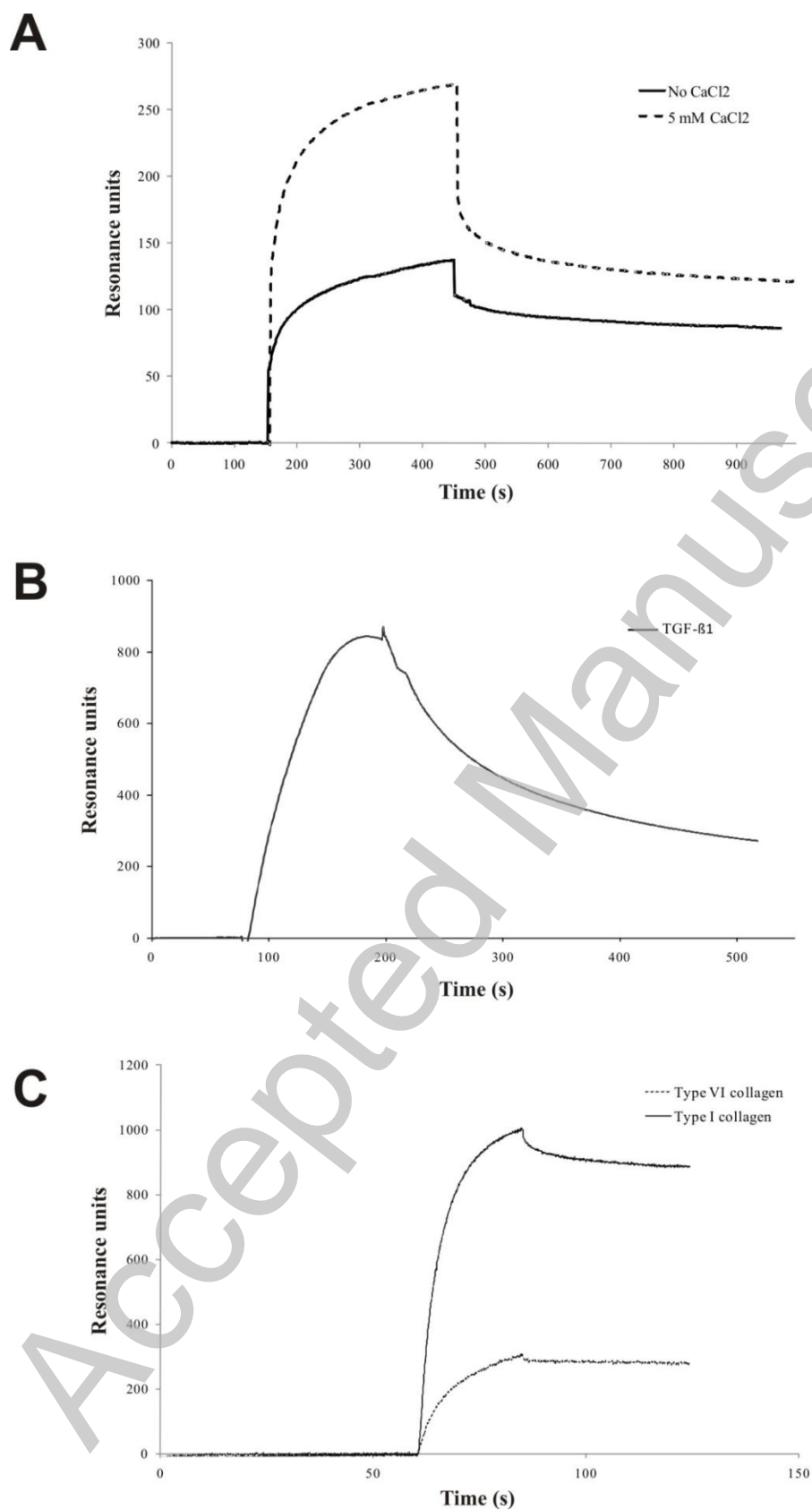


Figure 2

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

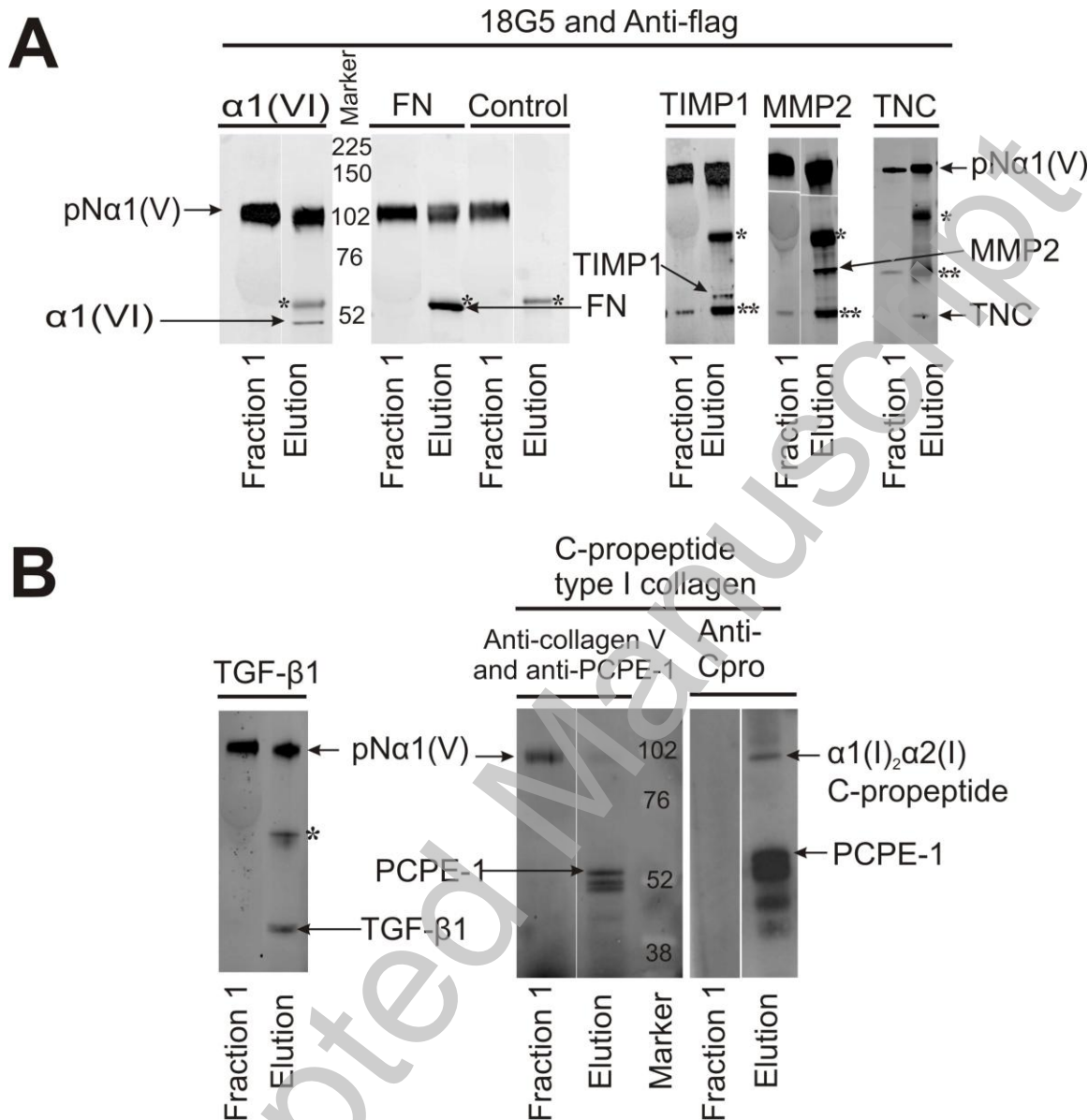


Figure 3

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

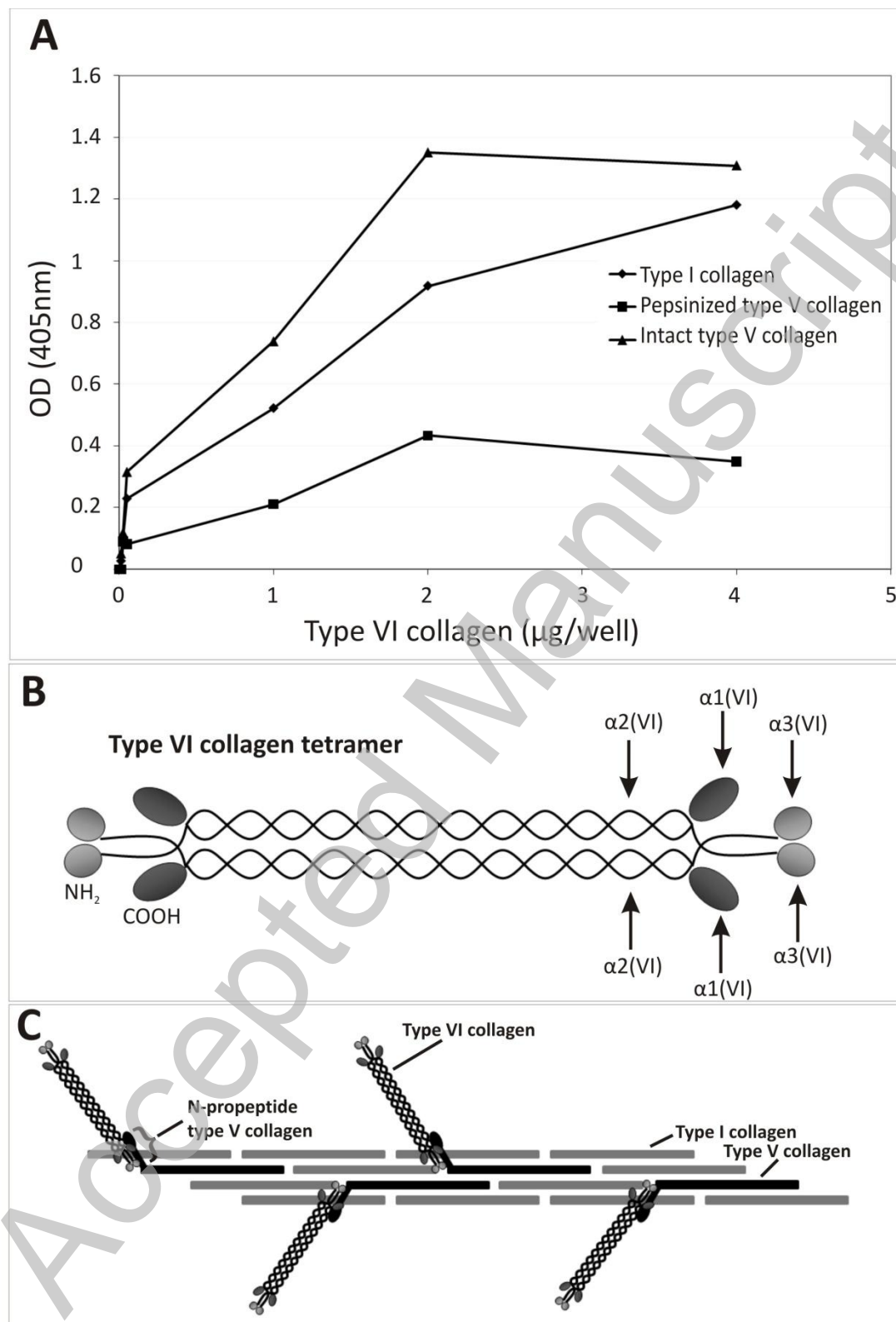


Figure 4

Identification of $\alpha 1(V)$ -N-propeptide interacting proteins

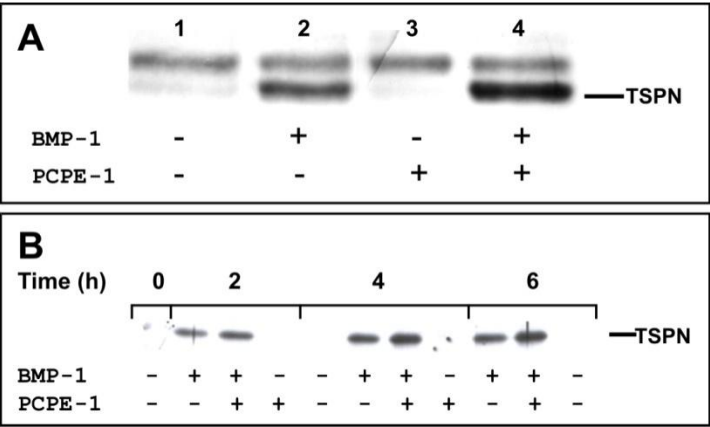


Figure 5