A Novel Marker of Tissue Junctions, Collagen XXII*

Manuel Koch‡‡, Joerg Schulze‡‡, Uwe Hansen¶, Todd Ashwoldt ¶, William J. Brunken‡‡, Robert E. Burgeson¶, Peter Bruckner¶¶, and Leena Bruckner-Tuderman §§

From the ‡‡Department of Dermatology, University of Freiburg, 79104 Freiburg, Germany, ¶¶Cutaneous Biology Research Center, Massachusetts General Hospital/ Harvard Medical School, Charlestown, Massachusetts 02129, ¶Center for Biochemistry, Medical Faculty University of Cologne, 50931 Cologne, Germany, §§Shriners Hospital for Children, Portland, Oregon 97201, ¶¶Anatomy and Cell Biology, Departments of Neuroscience, and Ophthalmology, Tufts University School of Medicine, Boston, Massachusetts 02111, ¶¶Department of Biochemistry, University of Muenster, 48129 Munster, Germany

Here we describe a novel specific component of tissue junctions, collagen XXII. It was first identified by screening an EST data base and subsequently expressed as a recombinant protein and characterized as an authentic tissue component. The COL22A1 gene on human chromosome 8q24.2 encodes a collagen that structurally belongs to the FACIT protein family (fibril-associated collagens with interrupted triple helices). Collagen XXII exhibits a striking restricted localization at tissue junctions such as the myotendinous junction in skeletal and heart muscle, the articular cartilage-synovial fluid junction, or the border between the anagen hair follicle and the dermis in the skin. It is deposited in the basement membrane zone of the myotendinous junction and the hair follicle and associated with the extracellular matrix in cartilage. In situ hybridization of myotendinous junctions revealed that muscle cells produce collagen XXII, and functional tests demonstrated that collagen XXII acts as a cell adhesion ligand for skin epithelial cells and fibroblasts. This novel gene product, collagen XXII, is the first specific extracellular matrix protein present only at tissue junctions.

Tissue integrity of all organs is critically dependent on suprabasemental aggregates containing collagens. The family of collagens is specified in man by 42 genes encoding polypeptides assembled into at least 28 distinct, trimeric collagens. Their functions are indirectly illustrated by a multitude of human diseases resulting from mutations in collagen genes. To date more than 1000 different mutations are known to cause “collagen diseases” in a wide spectrum of organ systems, including the skeletal system, ligaments and other soft connective tissues, the kidney, bone marrow, skin, eye, and as shown recently, the brain (1–3).

Depending on their occurrence in supramolecular assemblies and other structural features, collagens are subdivided into different classes, such as fibrillar, network-forming, beaded filament-forming, fibril-associated, or transmembrane collagens. The major fibrillar collagens often have a wide tissue distribution in mesenchymal connective tissues, such as bone, cartilage, tendons, or dermal connective tissue (1). However, other collagens, e.g. the network-forming basement membrane collagens, can exhibit a very limited tissue localization lining epithelia, endothelia, or muscle cells and separating them from the surrounding extracellular matrix. Such a restricted expression pattern is believed to indicate both a highly specialized mechanical role in maintaining integrity of a tissue compartment and a role in regulation of cellular functions (4).

The FACITs, fibril-associated collagens with interrupted triple helices, are quantitatively minor collagens that often copolymerize into suprastructure with the major collagens and mediate ligand interactions between the fibrils and their environment (5). Typically, these collagens contain triple helical as well as other functional protein modules, including VWA domains and fibronectin type III-like domains. VWA domains are found in a variety of proteins, e.g. the prototype von Willebrand factor, collagens, matrins, and integrins (for review, see Ref. 6). The general notion about the function of VWA domains is mediation of protein-protein interactions. For example, the classical collagen binding receptors, integrin α1β1 and α5β1, bind to their target via VWA domains (7). Although not yet proven, it is likely the VWA domains of FACIT collagens also bind to other proteins. Two new VWA-containing collagens, collagens XX and XXI, were recently identified by screening databases (8, 9). However, except for the gene structure and the predicted molecular domain structure, very little is known about these molecules, their distribution, or their functions in tissues.

Tissue junctions have critical functions in joining tissue compartments and in transmitting forces. However, because of a lack of specific markers for such junctions, their molecular and cellular composition and morphogenesis have remained elusive. The only exception may be the myotendinous junction

---

* This work was supported in part by an award from the Alexander von Humboldt Foundation sponsored by the German Federal Ministry of Education and Research (to M. K.), the German Research Council, Deutsche Forschungsgemeinschaft Grants Nr. Br 14757/1–KO 2247/1–1 (to L. B.-T. and M. K.), SFB 492/A2 (to P. B.), and SFB 492/A3 (to L. B.-T.), Public Health Service Grant R01 NS39503 (to R. E. B. and W. J. B.), and a Shiozaki award to the Cutaneous Biology Research Center Harvard Medical School (to R. E. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed: Dept. of Dermatology, University of Freiburg, Hauptstrasse 7, 79104 Freiburg, Germany. Tel.: 49-7612706716; Fax: 49-7612706936; E-mail: bruckner_tuderman@haut.ukl.uni-freiburg.de.

1 The abbreviations used are: FACIT, fibril-associated collagen with interrupted triple helices; VWA, von Willebrand factor A-like; TSPN, N-terminal thrombospondin-like domain; Coll XXII, collagen XXII; NC1 domain, corresponds to the N-terminal non-collagenous region of Col XXII; MTJ, myotendinous junction; kb, kilobase(s); RACE, rapid amplification of cDNA ends; pAb, polyclonal antibody.
(MTJ), which represents the link between muscles and tendons, the biology and pathology of which has been studied in more detail. To overcome the drastic forces at MTJs, the muscle increases the contact area by forming finger-like interdigitations of the basement membrane zone at the junction. Some molecules are enriched but not exclusively present at the MTJ, e.g. integrin α₁β₁, or tenasin (10–12). Consequently, mutations in integrin α₁β₁ cause muscular dystrophy (13, 14). Recent developmental studies have provided evidence for the requirement of communication between muscle and tendon for embryonic development. The non-embryonic origins of the muscle and tendon cells was shown to be the somites (12, 15); in muscle cell ablation experiments tendons were not formed in a muscle-less wing (16). However, except for the initial generation of the precursors, practically nothing is known about the signaling between the tissue compartments, which will lead to formation of the MTJ. In this study we identified and characterized a novel marker, collagen XXII (Col XXII), which exhibits a unique localization at tissue junctions in the muscle, tendons, heart, articular cartilage, and skin.

EXPERIMENTAL PROCEDURES

\textbf{cDNA Isolation}—A BLAST search (17) of the data base of expressed sequence tags (dbEST, Ref. 18) for homology with the C-terminal amino acid sequence of human collagen XII (accession number, NM_004370) yielded one clone (GenBank™ accession number AA299694) as a possible candidate for a novel collagen cDNA. The EST clone, which was about 1520 bp in length, was purchased from the American Type Culture Collection (ATCC) and sequenced in its entirety. The sequence revealed codons for 167 amino acid residues, and the 3′-untranslated region was 1.0 kb in length. From the sequence of the EST clone, nested primers were designed for 5′ RACE using a human placental cDNA Marathon Library (Clontech, Palo Alto, CA) as template, as previously described (19). For 5′ RACE, the Long Expand PCR kit (Roche Applied Science) was used for the PCR reaction. By performing in total three RACEs with primer sequences derived from each previous RACE, overlapping segments representing the full-length 6.4-kb mRNA were obtained. To confirm the nucleotide sequence and as a control for PCR-induced nucleotide substitutions, gene-specific primers were used to re amplify the entire cDNA on human cartilage/bone cDNA. A first strand cDNA synthesis kit (BD Biosciences Clontech) was used to synthesize cDNA from total RNA using random primers following the manufacturer’s protocol; PCR was used to generate overlapping clones complementary to the entire human a1 (XXII) collagen mRNA. Sequencing of all the PCR products obtained from the cDNA confirmed the nucleotide sequence of the human collagen XXII. The full-length cDNA of Col XXII is deposited in GenBank™ under accession number AF046780.

To generate a genetic relationship map of the VWA domain-containing proteins, the VWA of protein sequences were analyzed with the GrownTree program (GCG Sequence Analysis Software Package, Genetics Computer Group Inc.). The following sequences were used: matrilin1 (NP_000370, 35–221); matrilin2 (NP_005072, 51–237); matrilin3 (NP_002372, 72–263); matrilin4 (NP_003824, 22–224); collagen XII 1 (NP_004361, 1192–1376); collagen XII 2 (NP_004361, 2317–2501); collagen XIV 1 (CAAT2420, 1–185); collagen XIV 2 (XX, 151–353); collagen XXI (NP_110447, 31–212); collagen XXII (AAN03620, 32–218). The VWA were analyzed, and the Jukes-Cantor method was chosen to correct the distances for multiple substitutions at a single site; the tree was created with the unweighted pair group method using arithmetic averages (UPGMA) algorithm.

\textbf{Northern Blot Analysis}—A 1368-bp PCR product (AF046780, nucleotides 529–1896) was labeled with [32P]dCTP (PerkinElmer Life Sciences) using the rediprime DNA-labeling system (Amersham Biosciences). The Northern blot (Clontech) was prehybridized in 50% formamide, 5× saline/sodium phosphate/EDTA, 1× Denhardt’s, 1% SDS, 10% dextran sulfate, and 0.1 mg/ml salmon sperm DNA (Invitrogen) at 42 °C for 2 h. Without further purification the probe was denatured in the gel to a buffer plus 1× human Cot-1 DNA (Roche Applied Science) and 1:10 (v/v) sheared salmon testis DNA (Invitrogen) at 94 °C for 5 min, placed on ice, added to the blots, and hybridized for 20 h. Blots were washed 3 times in 2× SSC (1× SSC = 0.15 M NaCl and 0.015 M sodium citrate), 1% SDS at 42 °C, and 2 times in 0.1× SSC, 1% SDS at 42 °C. Blots were placed on BioMax MR film (Eastman Kodak Co.) with a BioMax TranScan-LE intensifying screen (Kodak) for 20 h at −80 °C.

\textbf{Immunoblotting}—The recombinant proteins were subjected to rotary shadowing using previously published instructions (Herculase DNA polymerase; Stratagene). The PCR product was purified on an agarose gel (Qiagen) and subcloned (rapid DNA ligation kit; Roche Diagnostics) into a modified PCEP-4 (gift from Ernst Poelsch) expression vector. For convenience, an His-tag followed by a thrombin cleavage site was included adjacent to the NheI site in the vector. The ligated DNA was transformed into TOP 10 cells (Invitrogen). Plasmids were isolated from the bacteria (Qiagen) and sequenced with gene-specific primers (Thermo Sequenase cycle sequencing kit; Amersham Biosciences). 293-EBNA cells (Invitrogen) were transformed (FuGENE; Roche Diagnostics) with the expression vector and selected after 2 days with puromycin (Sigma). Stably transfected 293-EBNA cells were pseudo-subcloned, and the highest protein-producing clones were expanded for large scale production. Two liters of supernatant from these cells were collected and supplemented with 1 mm Pefablock (Merck). After ammonium sulfate precipitation (45% saturation), the precipitate was collected by centrifugation, dialyzed against the binding buffer (200 mM NaCl, 20 mM Tris-HCl, pH 8), and applied onto a nickel-chelated Sepharose column (Amersham Biosciences). The elution was done by applying binding buffer containing increasing concentrations of imidazole (10–150 mm) to the column. In some cases, the His tag was digested with thrombin (isolated from bovine plasma; Sigma) according to the protocol from EMD Biosciences. The digested protein was again applied to a nickel-chelated Sepharose column to remove the His tag, and since the protein still weakly binds to the matrix, the fragment was again eluted with increasing imidazole concentration.

\textbf{Immunoblotting}—The human Col XXII NC1 protein was injected intradermally into a rabbit (R34) for antibody production following standard procedures (20). The R34 antiserum was passed over a protein A column (Amersham Biosciences) and eluted with triethylamine (Sigma). The neutralized eluate was affinity-purified by applying it to a human Col XXII NC1 protein column that was prepared by coupling the protein without His tag to activated CNBr-Sepharose. Bound antibodies were eluted with triethylamine and immediately neutralized (19).

\textbf{Immunodetection of Collagen XXII}—The above polyclonal antibody pAb R34 was used for immunofluorescence staining of cryosections of 1-day-old mice and adult mouse tissues. The incubation with the first antibody (1:2000) was done overnight at 4 °C followed by a 1-h incubation with fluorescein isothiocyanate- or rhodamine-coupled secondary antibody at room temperature. The staining was observed by immunofluorescence microscopy. For immunofluorescence microscopy the proteins were separated on SDS-PAGE with 4.5, 5.0, or 7.5% polyacrylamide under non-reduced or reduced conditions and transferred onto nitrocellulose. The affinity-purified antibodies were diluted 1:10,000 and incubated overnight followed by an incubation with horseradish peroxidase-linked anti-rabbit secondary antibody (Amersham Biosciences) for 2 h. The signals were visualized with chemiluminescence substrate Renaissance™ (PerkinElmer Life Sciences).

For immunoelectron microscopy with pAb R34, recombinant Col XXII was subjected to rotary shadowing using previously published methods (21). Native fibrils were isolated from cartilage, placed on grids, immunostained with pAb R34 and colloidal gold-labeled secondary antibodies, and analyzed with transmission electron microscopy as described (22). Immunoelectron microscopy with immunogold labeling on ultra-thin sections of skin, muscle, and articular surface was carried out as described previously (23).

\textbf{Tissue Extractions}—For analysis of Col XXII tissue form, adult mouse tissues were dissected, homogenized in Tris-buffered saline (1 g tissue/10 ml), and isolated sequentially at 4 °C with the following buffers; TBS for 30 min; 0.5 M NaCl, 50 mM Tris-HCl, pH 7.4, overnight; 2 M urea, 0.5 M NaCl, 50 mM Tris-HCl, pH 7.4, overnight. All buffers contained 10 mM EDTA, 1 mM N-ethylmaleimide, and 1 mM Pefabloc (Merck) as proteinase inhibitors. Between the extractions the tissue was centrifuged at 14,000 g at 4 °C for 30 min, and 50–100 μl of the supernatants were used for collagenase digestion. The tissue digestion with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion. The incubation with 40 units/ml highly purified collagenase digestion was targeted to collagenase digestion.
fied bacterial collagenase (Advanced Biofacturers Inc., Lynbrook, NY) was carried out in 50 μl of extraction buffer containing 5 mg CaCl₂ and 1 mg Pefabloc at 37 °C for 4 h (34). The reaction was stopped by adding EDTA to a final concentration of 20 mM.

**Cell Adhesion Assays**—Skin epithelial cells (HaCaT) and fibroblasts (WI-26) were cultivated under standard conditions. Multwell tissue culture plates (96 wells, Costar Corp., Faust, Germany) were coated with serial dilutions of Col XXII (0–165 mM) overnight at 4 °C. After saturation with 1% bovine serum albumin (fraction V, Sigma), the plates were immediately used for short term cell adhesion assays. After 30 min the floating cells were washed away, and the number of adherent cells were counted (24). All assays were done in triplicate.

**In Situ Hybridization**—In situ hybridization was performed as previously described (25) using cRNA probes generated from PCR products. The T7 polymerase recognition site was added to the reverse primers. For mouse collagen COL1A1 the region between nucleotides 3653 and 4242 (BC050014) was chosen as a probe. The mouse cDNA was obtained by comparison of the human Col XXII cDNA to the mouse genome. The following primers localized within the VWA and thrombospondin N-terminal-like domain (TSPN) domain region was used to amplify the probe from embryonic mouse cDNA (forward, 5′-GCCACCT-TCAACTCTCGCAGGGAGG-3′; reverse, 5′-CACAAAGGGACGCTCAGCTTGC-3′). Images shown here were processed using Adobe Photoshop (Releases 7.0); no enhancements other than contrast and brightness have been made to these images.

**RESULTS**

**Identification and Cloning of Collagen XXII**—Several novel partial cDNA sequences were identified in a human dbEST sequence data base search for clones containing Gly-X-Y triplets. One of these sequences was extended using rapid PCR amplification of cDNA ends. After three rounds of amplification, the full-length cDNA for Col XXII was obtained. It contains a predicted open reading frame of 1626 amino acids, including a putative signal peptide (27 amino acids) (Fig. 1; Releases 7.0). As predicted from the corresponding cDNA sequence, the mature protein is predicted to contain a long collagenous domain structure (9). The collagenous domain, however, is shorter in collagen XXI, in which the first collagenous stretch at the C-terminal end contains the imperfections (circled interrupted by several larger amino acid stretches and six smaller imperfections (circled)). The collagenous region is interrupted by several larger amino acid stretches and six smaller imperfections (circled) of the Gly-X-Y triplets.

**As judged by their domain organizations, the closest relative of Col XXII is human collagen XXI; it contains the same noncollagenous domain structure (9). The collagenous domain, however, is shorter in collagen XXI, in which the first collagenous stretch is missing (Fig. 2). Both molecules contain two cysteines on either side of the collagenous domain. The Collagen XXII VWA domain is 43% identical to the VWA domains of human collagen XXI, 37% in collagen XII and 36% in Matrilin-1 (Fig. 3). The TSPN domain, typically located between the noncollagenous domain and the collagenous domain, is a common feature of proteins of the FACIT subfamily. Comparing the TSPN domain of Col XXII to those in other collagens indicated a close relation to collagen XII and XXI. However, the overall homologies between the different TSPN domains were too low for the construction of a statistically relevant phylogenetic tree.

**The Collagen XXII Gene COL22A1**—Comparison of the COL22A1 gene encoding collagen XXII with sequences in the genome databases revealed an orthologue for collagen XXII in mouse, zebra fish, and puffer fish. COL22A1 is present on human chromosome 8q24.2 and spans 326 kb (NT_028251).
removal of the His tag by cleavage with thrombin, the protein was used for immunization. The affinity-purified rabbit antibody pAb R34 efficiently recognized the NC1 domain of Col XXII in immunoblots. The calculated molecular mass of this domain is 53 kDa; however, the recombinant protein migrates with an electrophoretic mobility corresponding to an apparent mass of 70 kDa (Fig. 4A, lane 4). Removal of the His tag had no effect on the ability of the antibodies to recognize Col XXII (Fig. 4A, lane 4). Importantly, despite structural similarities between this domain in collagens XXI and XXII, no cross-reactivity with collagen XXI was seen (Fig. 4A, lane 3).

Recombinant Full-length Col XXII—Recombinant human Col XXII was expressed in 293-EBNA cells and affinity-purified using the His tag (Fig. 4B, lanes 1 and 3). Under non-reducing conditions, monomeric, dimeric, and trimeric molecules were detected, indicating partial inter-chain disulfide bonding (Fig. 4B, lane 2). The monomeric polypeptide chains of full-length Col XXII had an electrophoretic mobility consistent with a mass of about 200 kDa (Fig. 4B); in some cases a smaller product of about 70 kDa, corresponding to the NC1 domain, was observed. The full-length molecule was collagenase-sensitive, and digestion yielded the collagenase-resistant NC1 domain (Fig. 4B, lane 4).

Visualization of recombinant Col XXII by transmission electron microscopy after rotary shadowing indicated that the recombinant molecule formed structures resembling FACIT collagens (Fig. 5, A–C). Full-length Col XXII has a thin rod-like structure that corresponds to the collagenous domain, with a contour length of about 301 ± 15 nm. The rod contains several kinks representing the interruptions of the collagenous domain. At the N-terminal region the globules correspond to the VWA and TSPN domains (Fig. 5, A and D). From these images the flexible structure of Col XXII becomes evident.

Col XXII in Tissue Extracts—Col XXII was easily extracted with a high salt buffer from muscle and, less efficiently, from skin (Fig. 6A, lanes 1 and 2). Very little additional protein was released by subsequent urea extraction of either tissue (Fig. 6A, lanes 4 and 5). In addition to the 200-kDa Col XXII band, smaller bands of 110–120 and 70 kDa were visible in an immunoblot. Indeed, Col XXII proved to be very sensitive to proteolysis. Despite rapid protein isolation methods and use of potent proteinase inhibitor cocktails during extraction, a single band was not seen in tissue extracts. Similarly, if the recombinant full-length protein was stored for some time at 4 °C, a comparable fragmentation pattern was obtained (not shown). Collagenase digestion of tissue extracts abolished the 200-kDa and the 160-kDa bands, and immunoblotting with pAb34 demonstrated that the 70-kDa band corresponded to the NC1 domain (Fig. 6A, lane 3). These observations confirmed that in
A three globular domains represent the non-collagenous interruptions. The N terminus contains a rod-like structure representing the collagenous region, and the kinks represent single globules.

In the lung, liver, or kidney, Col XXII mRNA was detected only in the muscle cells. In heart and skeletal muscle but not in other organs such as brain, placenta, and kidney, only in the muscle cells. In high salt, four distinct bands were detectable in muscle extracts. Treatment of the extract shown in lane 2 with collagenase abolished the bands with higher molecular weight, indicating that they had a collag-

Fig. 5. Rotary shadowing images of purified recombinant Col XXII. A, three examples of a full-length molecule are shown. The thin rod-like structure represents the collagenous region, and the kinks represent the non-collagenous interruptions. The N terminus contains three globular domains (arrow). D, recombinant NC1 domain forms single globules. Bars, 50 nm.

Fig. 6. Expression of Col XXII in tissues. A, immunoblot with pAb R34 of proteins extracted from mouse skin and muscle using high salt (lanes 1–3) and urea (lanes 4 and 5). The skin extracts contained relatively small amounts of Col XXII (lane 1). In contrast, in high salt extract of muscle, four distinct bands were detectable (lane 3). Treatment of the extract shown in lane 2 with collagenase abolished the bands with higher molecular weight, indicating that they had a collagenous structure. Only the NC1 domain resisted the protease digestion (lane 3). As seen here, the NC1 domain occasionally migrated as a double band, probably due to differential glycosylation of the molecules (lanes 2 and 3). Additional extraction of the tissues with 7 M urea did not result in recovery of more Col XXII from the skin (lane 4) or muscle (lane 5). B, Northern blot of Col XXII in different tissues. Only a single mRNA of 6.4 kb was detected on a human tissue blot. A strong signal was found in the heart and skeletal muscle but not in other organs such as brain, placenta, lung, liver, or kidney. In A molecular weight standards on the left are shown in kDa; in B the RNA size standards are in kb.

The tissue extracts, partial proteolytic degradation of the collagenous stretch of the molecule had taken place.

Tissue Distribution of Col XXII mRNA—Northern blot analysis of human tissues showed Col XXII mRNA to be highly expressed in skeletal muscle and heart (Fig. 6B). A single band of about 6.4 kb was detected, indicating no alternative splice variants. The size of the COL22A1 mRNA corresponds well to the length of the cloned cDNA. Semiquantitative RT-PCR on mouse tissues revealed additional signals in cartilage, skin, and keratinocytes and in the eye (not shown). In contrast, almost no signal was obtained from neuronal tissues or other organs such as bone, liver, kidney, or lung.

Col XXII mRNA Is Expressed in Muscle Cells, Not in Fibroblasts—By in situ hybridization, COL22A1 mRNA was detected exclusively in muscle cells at the muscle attachment sites to tendon elements and ribs (Fig. 7, A and B). Dense alkaline phosphatase reaction products were observed only in the muscle cells closest to the rib or aponeurosis. No signal was detected in muscle fibers at any other location in the muscle. For comparison, a collagen I probe was used to identify fibroblasts at the tendinous sheet zone. As shown in Fig. 7, C and D, Col XXII mRNA is absent from the region positive for collagen I message. Controls with antisense probes remained negative (data not shown).

Tissue Distribution of Col XXII Protein—Tissue distribution of Col XXII was determined by indirect immunofluorescence in postnatal mouse tissues. A striking observation was that Col XXII is expressed only at sites of tissue junctions in muscle, cartilage, heart, and skin (Fig. 8). Col XXII was found juxtaposed to tendon insertion sites, i.e., in tendinous sheets, so called aponeurosis, which separate muscle compartments, express Col XXII mRNA. C, for comparison, collagen I mRNA is mainly synthesized by fibroblasts of the aponeurosis between the muscles. D, hybridization of a parallel section of C for Col XXII mRNA shows that Col XXII is not synthesized by the tendon fibroblasts (asterisks) but can be detected only in the muscle cells.

In articular cartilage a small narrow band of Col XXII was detected at the cartilage surface facing the synovial fluid (Fig. 8E). Similarly, ultrathin sections surface-labeled for Col XXII exhibited a narrow positive zone of gold particles close to the articular surface (Fig. 8D). Because Col XXII belongs to the FACITs, its association with collagen-containing fibrils was examined by immunoelectron microscopy of fragments of na-
tive cartilage fibrils isolated from tissue homogenates (27). Gold particles representing Col XXII were not localized to the large collagen-containing fibrils but were found in the filamentous extrafibrillar material surrounding the fibrils, such as fibrillin (Fig. 9, E and F).

In the skin, Col XXII was found in a striking layer surrounding the lower third of anagen hair follicles (Fig. 8G). In cross-sections, for example of the mouse tail, the staining was reminiscent of thin ring-like structures embedded in the dermal matrix (Fig. 8H). No signal was detected in the interfollicular epidermis or in the uppermost region of the hair follicle. However, a positive staining was consistently observed in the sebaceous glands associated with the hair follicles. At the lower region of the hair follicle, the staining coincided with the presence of myofibroblast (Fig. 8H). Immunoelectron microscopy showed the gold particles to be in and around the lamina densa of the follicular basement membrane, which was strongly invaginated with cellular protrusions and closely interconnected with thin cross-banded fibrils in parallel orientation with the follicle wall (Fig. 9C).

Cell Attachment to Col XXII—Because collagens are well known ligands for cell surface receptors, the cell binding properties of Col XXII were assessed. Integrins $\alpha_1\beta_1$ and $\alpha_2\beta_1$ are the classic “collagen receptors”; therefore, two cell types expressing these integrins, HACAT keratinocytes and WI-26 lung fibroblasts, were tested for binding to Col XXII. Both cells bound to Col XXII in a concentration-dependent manner, and saturation was reached already at the coating concentration of $10\mu g/ml$ (Fig. 10). Binding efficiency of the two cell types was clearly different; WI26, which expresses both $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins, bound more efficiently than HACAT cells, which express only $\alpha_2\beta_1$ integrins.

**DISCUSSION**

Most members of the collagen protein family are unique to vertebrates, and their geneses are closely linked to the evolution of bones, tendons, vasculature, and organs. Because not all collagens are deposited in tissues in great abundance, some family members have remained undiscovered. Recently, with the support of the human genome sequencing project and the expressed single sequence tag database, several new collagens have been traced (28, 29). Here we identified a short cDNA clone of Col XXII, a novel collagen, from the dbEST database. After definition of the full-length cDNA by RACE amplification, interpretation of the translated cDNA revealed the existence of a novel FACIT collagen. In addition to human, putative EST
were visualized with a colloidal gold-labeled second antibody. Thin sections were incubated with pAb R34, and the bound antibodies allowed by a short C-terminal collagenous stretch. The domain organization of the two molecules is very similar except that Col XXI lacks a large segment of the collagenous domain (Fig. 2; Ref. 9). The C-terminal collagenous domains of FACITs are believed to interact directly or indirectly with collagen-containing fibrils in tissues (30). However, our electron microscopic studies indicate that Col XXII is not directly associated with collagen-containing fibrils. Rather, based on the observations on native fibril extracts from articular cartilage, it seems to interact with components of microfibrils, such as fibrillins or collagens VI. A similar situation prevails for collagen XVI, another FACIT protein, which was recently shown to be associated with fibrillin (22).

The structure of the recombinantly expressed Col XXII, as seen in the electron microscope after rotary shadowing (Fig. 5), fits well with the calculated length of the molecule. With the total number of 1045 amino acid residues in the collagenous domains and the 0.289-nm rise per residue in a triple-helical conformation, the predicted length of the triple-helical rod is 302 nm. This is in excellent agreement with the observed 304 ± 15-nm tail. The sequence of Col XXII contains four major and several minor interruptions of the collagenous domain, a fact that explains the flexibility of this molecule.

The tissue form of Col XXII was identified by immunoblotting of mouse tissues with an antibody against the recombinant NC-1 domain of human Col XXII (antibody pAb R34). After Northern blot and RT-PCR analysis of different tissues and cells revealed strong expression of COL22A1 in muscle and heart (and weaker in cartilage, skin, and eye), muscle and skin extracts were analyzed first. In immunoblots of both tissues, pAb R34 recognized a 200-kDa band that corresponds to the authentic full-length Coll XXII molecule. In muscle extracts, additional bands were seen. These most likely represent proteolytic cleavage products generated during extraction and protein chemical analysis. The non-collagenous NC-1 domain was stable, and the cleavage occurred within the collagenous regions. This is a rather common phenomenon for large secreted proteins, such as collagens and laminins (31, 32). However, the fact that Col XXII and Col XXI are highly homologous led us to exclude cross-reactivity carefully. First, immunoblots overloaded with the recombinant NC-1 domain of human Col XXI did not exhibit a signal with the antibody pAb R34. Second, an intensive search for the COL21A1 gene in the mouse genome databases did not yield evidence for the existence of murine collagen XXI. Thus, the tissue form of monomeric collagen XXII corresponds to that of the recombinant molecule visualized by rotary shadowing electron microscopy. However, the functional suprastructure form of Col XXII in situ remains unknown at present.

The tissue distribution pattern of Col XXII is unique. By immunofluorescence staining, the protein was localized to specific tissue junctions. In the skin, a sheet-like structure sur-
ranging the hair follicle contains Col XXII. This sheet is Col XXII-positive in areas of the lower follicle, where myofibroblasts line the outer surface of the hair follicle along the junction between the follicle and the dermis (33). In the joints, the surface of the articular basement membrane, i.e. the junction between cartilage and synovial fluid, was labeled with Col XXII antibodies within a very thin, confined band. At the ultrastructural level, this region contains highly organized thin collagen-containing fibrils and other suprastructural elements. Our experiments with extraction of native fibrils from articular cartilage demonstrated that Col XXII is not associated with the classic cross-banded “collagen fibrils” but, rather, with microfibrils. In arthritic human joints, Col XXII is still detectable with immunofluorescence staining, but the staining pattern is broadened and fuzzier.2 Future studies will show how Col XXII is integrated into functional suprastructures of the articular surface and around the hair follicle and which the cellular origin of this protein is.

In muscle, heart, and ciliary body, Col XXII was localized at the insertion sites of tendons or zonula fibrils into the muscle. The scaffolds, which are necessary for the integrity of tissues and the transmission of locomotive forces, are formed by polymeric protein structures, in which collagens play an integral role (34). Of all the sites in which Col XXII was found, the MTJ is best studied.

MTJ are crucial elements in the transmission of mechanical force from the muscle via tendons to the skeletal elements. Several molecules have been identified at this site. Tenascin, an oligomeric extracellular matrix protein, was one of the first markers identified (11), but its function at this site still remains unknown. Other components, such as $\alpha_1\beta_1$ integrin, laminin 2, and the dystrophin glycoprotein complex are pivotal, since mutations in their genes lead to pathologic changes of the articular surface, containing fibrils and other suprastructural elements. Our experiments with extraction of native fibrils from articular cartilage and synovial fluid, was labeled with Col XXII antibodies within a very thin, confined band. At the ultrastructural level, this region contains highly organized thin collagen-containing fibrils and other suprastructural elements. Our experiments with extraction of native fibrils from articular cartilage demonstrated that Col XXII is not associated with the classic cross-banded “collagen fibrils” but, rather, with microfibrils. In arthritic human joints, Col XXII is still detectable with immunofluorescence staining, but the staining pattern is broadened and fuzzier.2 Future studies will show how Col XXII is integrated into functional suprastructures of the articular surface and around the hair follicle and which the cellular origin of this protein is.

In muscle, heart, and ciliary body, Col XXII was localized at the insertion sites of tendons or zonula fibrils into the muscle. The scaffolds, which are necessary for the integrity of tissues and the transmission of locomotive forces, are formed by polymeric protein structures, in which collagens play an integral role (34). Of all the sites in which Col XXII was found, the MTJ is best studied.

MTJ are crucial elements in the transmission of mechanical force from the muscle via tendons to the skeletal elements. Several molecules have been identified at this site. Tenascin, an oligomeric extracellular matrix protein, was one of the first markers identified (11), but its function at this site still remains unknown. Other components, such as $\alpha_1\beta_1$ integrin, laminin 2, and the dystrophin glycoprotein complex are pivotal, since mutations in their genes lead to pathologic changes of the articular surface, containing fibrils and other suprastructural elements. Our experiments with extraction of native fibrils from articular cartilage and synovial fluid, was labeled with Col XXII antibodies within a very thin, confined band. At the ultrastructural level, this region contains highly organized thin collagen-containing fibrils and other suprastructural elements. Our experiments with extraction of native fibrils from articular cartilage demonstrated that Col XXII is not associated with the classic cross-banded “collagen fibrils” but, rather, with microfibrils. In arthritic human joints, Col XXII is still detectable with immunofluorescence staining, but the staining pattern is broadened and fuzzier.2 Future studies will show how Col XXII is integrated into functional suprastructures of the articular surface and around the hair follicle and which the cellular origin of this protein is.