Three Novel Collagen VI Chains with High Homology to the α3 Chain*□S

Sudheer Kumar Gara[‡1]**, Paolo Grumati**[§]**, Anna Urciuolo**[§]**, Paolo Bonaldo**[§]**, Birgit Kobbe**[‡]**, Manuel Koch**[‡¶‖]**,
Mats Paulsson**[‡¶]**, and Raimund Wagener**[‡2]

*From the ‡Center for Biochemistry, ¶Center for Molecular Medicine, and ‖Department of Dermatology, Medical Faculty,
University of Cologne, D-50931 Cologne, Germany and the §Department of Histology, Microbiology and Medical Biotechnologies,
University of Padova, 35121 Padova, Italy*

Here we describe three novel collagen VI chains, α4, α5, and α6. The corresponding genes are arranged in tandem on mouse chromosome 9. The new chains structurally resemble the collagen VI α3 chain. Each chain consists of seven von Willebrand factor A domains followed by a collagenous domain, two C-terminal von Willebrand factor A domains, and a unique domain. In addition, the collagen VI α4 chain carries a Kunitz domain at the C terminus, whereas the collagen VI α5 chain contains an additional von Willebrand factor A domain and a unique domain. The size of the collagenous domains and the position of the structurally important cysteine residues within these domains are identical between the collagen VI α3, α4, α5, and α6 chains. In mouse, the new chains are found in or close to basement membranes. Collagen VI α1 chain-deficient mice lack expression of the new collagen VI chains implicating that the new chains may substitute for the α3 chain, probably forming α1α2α4, α1α2α5, or α1α2α6 heterotrimers. Due to a large scale pericentric inversion, the human COL6A4 gene on chromosome 3 was broken into two pieces and became a non-processed pseudogene. Recently COL6A5 was linked to atopic dermatitis and designated COL29A1. The identification of novel collagen VI chains carries implications for the etiology of atopic dermatitis as well as Bethlem myopathy and Ullrich congenital muscular dystrophy.

Members of the collagen protein superfamily play important roles in maintaining extracellular matrix structure and func-tion. To date 28 family members are known (1, 2), among which the fibril-forming collagens and the FACIT collagens form large subgroups. In addition, several collagens exist that have highly specific functions. Among these, collagen VI forms a distinct network of microfibrils in most connective tissues. Electron microscopy revealed a beaded filament structure of the microfibrils (3). The α1, α2, and α3 chains of collagen VI form heterotrimeric monomers that already intracellularly assemble to dimers and tetramers (4, 5). After secretion, fila-ments are formed by end to end interactions of the preas-sembled tetramers.

The three previously known collagen VI chains contain a relatively short collagenous domain of about 335 residues together with VWA[3] domains, which are the characteristic non-collagenous domains of collagen VI. A common feature of VWA domains is their involvement in the formation of multi-protein complexes (6). Whereas all three collagen VI chains contain two C-terminal VWA domains, the α1 and α2 chains carry only one and the α3 chain ten VWA domains at the N terminus (7, 8). In addition, the α3 chain contains a unique domain with similarities to salivary gland proteins, a fibronec-tin type III repeat, and a bovine pancreatic trypsin inhibitor/ Kunitz family of serine protease inhibitor domain (Kunitz domain) at the C terminus (8). It was suggested that the VWA domains play a role in the assembly of collagen VI (9–11). How-ever, recently the analysis of lysyl hydroxylase 3-deficient mouse embryos indicated that also the loss of potentially glyco-sylated hydroxylysine residues prevents the intracellular for-mation of collagen VI tetramers and leads to impaired secretion of collagen VI (12).

It has been shown that collagen VI interacts with several other extracellular matrix components, including collagen I (13), II (14), and XIV (15), perlecan (16), and the microfibril-associated glycoprotein MAGP1 (17). The N-terminal globular domains of the collagen VI molecules bind the small leucine-rich repeat proteoglycans decorin and biglycan, which in turn interact with matrilins, mediating contacts to further binding partners (18).

Studies on collagen VI have often focused on its function in skeletal muscle because of the patient phenotypes, Beth-

lem myopathy, and Ullrich congenital muscular dystrophy, observed when the $\alpha1$, $\alpha2$, or $\alpha3$ chain carries a mutation (for review see Ref. 19). In mice where the gene coding for the collagen VI $\alpha1$ chain has been inactivated also the $\alpha2$ and $\alpha3$ chains are not secreted, showing that a heterotrimeric assembly is required (20). The mice show a muscular weakness and histological signs of muscle fiber necrosis. Recent studies indicate that the myopathy is due to a mitochondrial dysfunction (21, 22). A possible explanation could be a decreased integrin-mediated signaling from collagen VI to the cells (23), but details of the downstream events are still not known.

Here we describe three new collagen VI chains that have the potential to replace the collagen VI $\alpha3$ chain in collagen VI assemblies and thereby to increase the structural and functional versatility of collagen VI.

MATERIALS AND METHODS

RT-PCR—RT-PCR was used to clone the mouse and human collagen VI cDNAs. Primers were designed according to EST and genomic sequences that are deposited in the data bases (supplemental Table 1). To prevent mutations in the RT-PCR we used the Expand high fidelity PCR system (Roche Applied Science). The cDNAs for the $\alpha4$ chain were amplified from mRNA isolated from adult mouse uterus and newborn mouse brain, and cDNAs for the $\alpha5$ and $\alpha6$ chains were amplified from mRNA from newborn mouse lung. The human cDNAs for the $\alpha5$ chain were cloned from mRNA prepared from HT1080 or HEK293-EBNA cells, and the cDNAs for the $\alpha6$ chain were cloned from mRNA prepared from fetal brain using the primer pairs indicated in supplemental Table 1.

Northern Blot Analysis—Total RNA was extracted from various tissues of newborn and adult C57BL/6J mice by the guanidinium-thiocyanate method. mRNA was prepared by using the Oligotex® mRNA Mini Kit (Qiagen). Aliquots were electrophoresed on a 0.8% denaturing agarose-formaldehyde gel, blotted, and hybridized with digoxigenin-labeled RNA probes. The conditions in the last two wash steps were: $0.1 \times$ SSC, 0.1% SDS at 68 °C for 15 min each. The blots were developed using CDP-Star (Roche) according to the manufacturer's instructions.

Bioinformatic Analysis—The non-redundant NCBI genomic data bases for mouse (Build 37.1) and human (Build 36.2) were scanned for new genes using collagen and matrilin sequences as queries. The exon-intron boundaries of each of the new genes were carefully interpreted using the NCBI Evidence Viewer together with the cloned cDNA sequences. The potential signal peptide and domain structure of each protein was predicted by SignalP v3.1 and SMART, respectively. However, the N1 domain of the $\alpha5$ chain was manually assigned based on sequence signature motifs because none of the available domain prediction programs could locate it. Multiple sequence alignments were performed using CLUSTAL X (v1.81) and figures were prepared with the BOXSHADE v3.2 program. The protein sequence identities of the new chains were calculated using BOXSHADE. The

phylogenetic analysis was done by protein distance and protein parsimony as described in PHYLIP v3.66.

Expression and Purification of Recombinant N-terminal Collagen VI $\alpha4$, $\alpha5$, and $\alpha6$ Chain VWA Domains—cDNA constructs were generated by RT-PCR on mRNA. For the collagen VI $\alpha4$, $\alpha5$, and $\alpha6$ chains, the domains N3–N6, N3, and N1–N7 were chosen, respectively. Suitable primers introduced 5′-terminal NheI and 3′-terminal BamHI, BglII, or XhoI restriction sites (supplemental Table 1). The amplified PCR products were inserted into a modified pCEP-Pu vector (16) containing an N-terminal BM-40 signal peptide and a C-terminal His$_8$-tag or a C-terminal tandem strepII-tag (17) downstream of the restriction sites. The recombinant plasmids were introduced into HEK293-EBNA cells (Invitrogen) using FuGENE 6 transfection reagents (Roche). The cells were selected with puromycin (1 μg/ml), and the His$_8$-tagged protein-producing cells were transferred to serum-free medium for harvest of the recombinant protein. The C-terminal tandem strepII-tagged protein was directly purified from serum-containing cell culture medium. After filtration and centrifugation (1 h, $10{,}000 \times g$), the cell culture supernatants were applied either to a streptactin column (1.5 ml, IBA GmbH) and eluted with 2.5 mM desthiobiotin, 10 mM Tris-HCl, pH 8.0, or to a TALON metal affinity column (Clontech) and eluted following the supplier's protocol.

Preparation of Antibodies against the New Collagen Chains—The purified recombinant collagen VI fragments were used to immunize rabbits and guinea pigs. The antisera obtained were purified by affinity chromatography on a column with antigen coupled to CNBr-activated Sepharose (GE Healthcare). The specific antibodies were eluted with 0.1 M glycine, pH 2.5, and the eluate was neutralized with 1 M Tris-HCl, pH 8.8. The antiserum raised against the domains N1–N7 of the collagen VI $\alpha6$ chain were affinity-purified on a column coupled with the collagen VI $\alpha6$ chain N2–N6 domains to prevent cross-reactivity due to the highly identical N7 domains of collagen VI $\alpha5$ and $\alpha6$. The lack of extensive cross-reactivity between the new chains was demonstrated by ELISA.

Immunohistochemistry—Immunohistochemistry was performed on frozen embedded sections of adult wild type and collagen VI $\alpha1$ chain-deficient mice (20). The frozen sections were preincubated in ice-cold methanol for 2 min, blocked for 1 h with 5% normal goat serum in phosphate-buffered saline containing 0.2% Tween 20, and incubated with the primary antibodies overnight at 4 °C followed by AlexaFluor 488-conjugated goat anti-rabbit IgG (Molecular Probes), AlexaFluor 546-conjugated goat anti-rabbit IgG (Molecular Probes), or AlexaFluor 488-conjugated goat anti-guinea pig IgG (Molecular Probes). Collagen VI $\alpha1$, $\alpha2$, and $\alpha3$ chains were detected using a polyclonal antibody (AB7821, Chemicon). A polyclonal antibody against the human native laminin-332 (24) was kindly given by R. E. Burgeson.

Preparation of Muscle Extracts—Frozen mouse skeletal muscle was pulverized by pestle and mortar and lysed with a solution containing 50 mM Tris, pH 7.5, 150 mM NaCl, 10 mM MgCl$_2$, 0.5 mM dithiothreitol, 1 mM EDTA, 10% glycerol, 2% SDS, 1% Triton X-100, 1 mM phenylmethylsulfonyl fluoride, 1
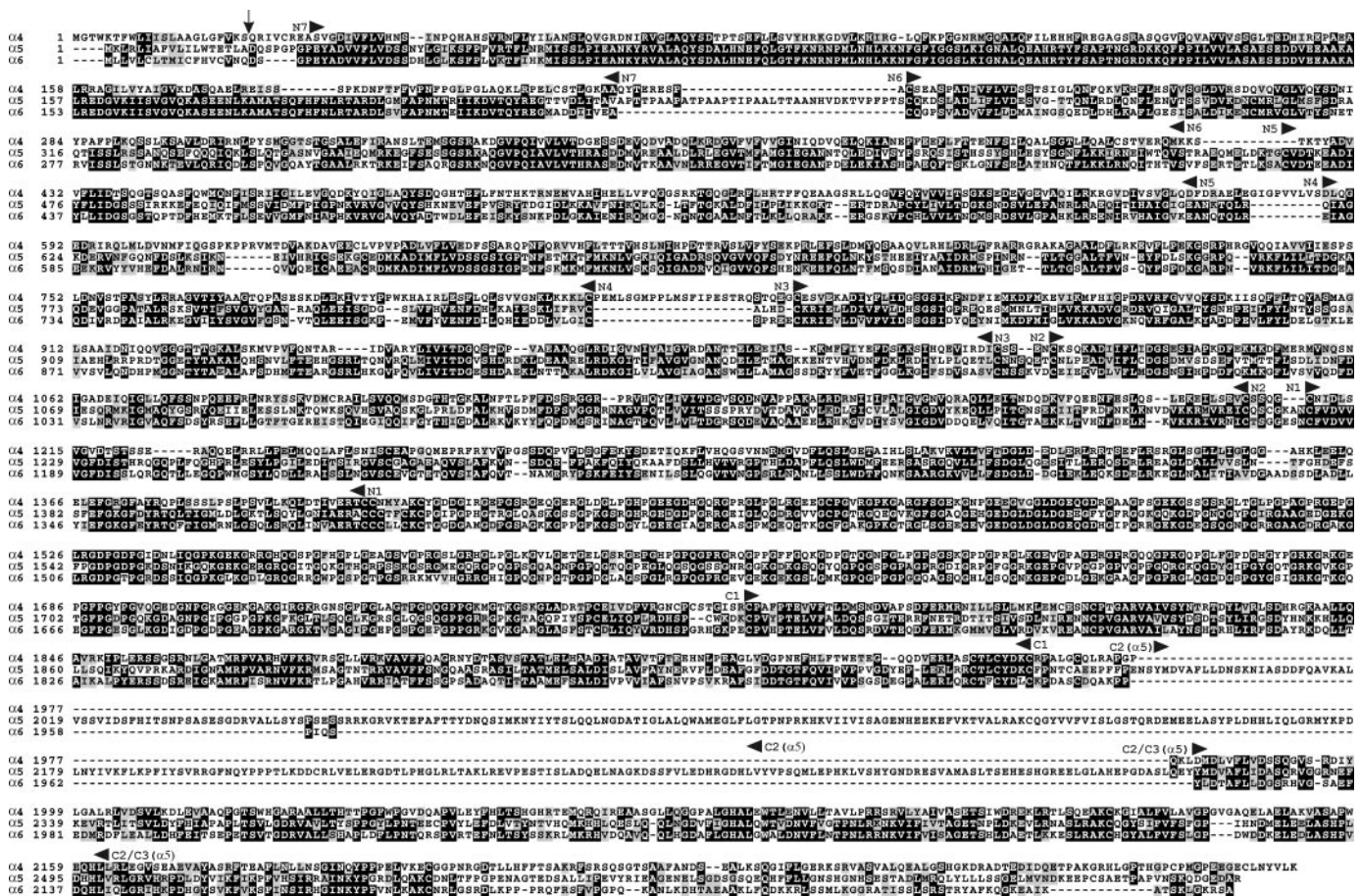
FIGURE 1. **Alignment of amino acid sequences of murine collagen VI α4, α5, and α6 chains.** The amino acid sequences were deduced from the cDNA sequences deposited in the data base under accession numbers AM231151–AM231153, AM748256–AM748258, and AM748259–AM748262, respectively. The *arrow* marks the potential signal peptide cleavage sites. *Arrowheads* indicate the boundaries of the domains depicted in Fig. 2.

mM sodium orthovanadate, 5 mM sodium fluoride, 3 mM glycerol 2-phosphate, and protease inhibitors (Complete, Roche). Proteins were solubilized by heating at 70 °C for 10 min. Samples were clarified by centrifugation at 4 °C.

Purification of Collagen VI—Native collagen VI was purified from newborn mice. Proteins were extracted by urea treatment, and collagen VI was isolated by molecular sieve column chromatography as described previously (25).

Gel Electrophoresis and Immunoblot—Samples were reduced with 5% β-mercaptoethanol and subjected to SDS-PAGE on 4–12% (w/v) gradient polyacrylamide gels. Proteins were electrophoretically transferred to Immobilon-P transfer membrane (Millipore). Collagen VI α1, α2, and α3 chains were detected using the 70-XR95 polyclonal antibody (Fitzgerald Industries International). The collagen VI α1 chain was detected using a polyclonal antibody recognizing the human α1(VI) chain (H-200, Santa Cruz Biotechnology). The new collagen VI α4, α5, and α6 chains were detected using the affinity-purified antibodies described above. As a loading control, an antibody against glyceraldehyde-3-phosphate dehydrogenase was used (MAB374, Chemicon). Secondary antibodies conjugated with horseradish peroxidase were used, and bands were detected by chemiluminescence (SuperSignal West Pico, Pierce).

RESULTS

Cloning of cDNAs Coding for Three New Mouse Collagen VI Chains—In a screen of the genomic data base with collagen and matrilin sequences as queries, three genes were identified in the mouse genome that code for new VWA domain-containing collagens. Because of their homology to the α3 chain of collagen VI and their arrangement in the genome, these were designated as the α4, α5, and α6 chains of collagen VI. The corresponding cDNAs were cloned as overlapping partial clones by RT-PCR, using primers deduced from the genomic sequence, and sequenced. The cloned mouse α4 cDNA of 7084 bp (accession numbers AM231151–AM231153) contains an open reading frame of 6927 bp, encoding a protein consisting of 2309 amino acid residues preceded by a signal peptide of 22 residues, as predicted by a method using neural networks or hidden Markov models, respectively (13). The mature secreted protein has a calculated M_r of 248,389 (Fig. 1). At least nine EST clones exist that extend 207 bp in the 3′ direction and contain an ATTAAA polyadenylation signal at their 3′-ends. In addition, a partial RIKEN cDNA clone (AK159050) extends 1219 bp and also contains an ATTAAA polyadenylation signal at its 3′-end, indicating the presence of different 3′-UTRs.

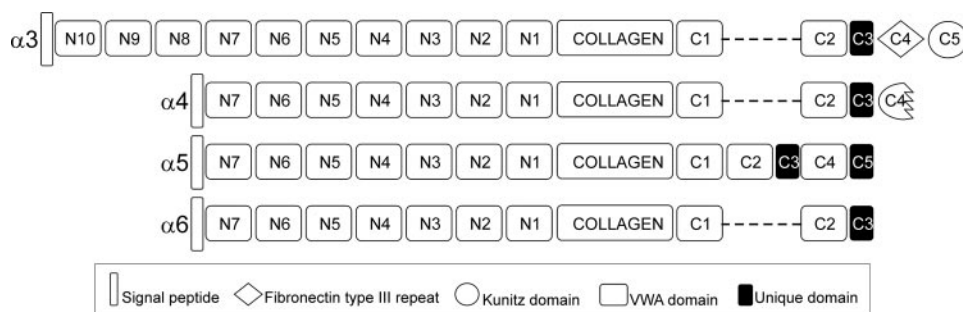The cloned mouse α5 chain cDNA of 8298 bp (accession numbers AM748256–AM748258) contains an open reading

FIGURE 2. **Domain structures of the new collagen VI chains compared with the collagen VI α3 chain.** Shown are the VWA domain and the Kunitz family of serine protease inhibitors domain (Kunitz domain). The numbering of the domains is according to Chu *et al.* (8). The *dashed lines* in the collagen VI α3, α4, and α6 chains indicate the lack of duplication of the second VWA domain and the unique domain found in the collagen VI α5 chain.

frame of 7920 bp, encoding a protein consisting of 2640 amino acid residues preceded by a signal peptide of 18 residues (13). The mature secreted protein has a calculated M_r of 287,502 (Fig. 1). A partial RIKEN clone (AK134435) extends 751 bp at the 3′-end but does not contain a polyadenylation signal.

The cloned mouse collagen VI α6 chain cDNA of 7097 bp (accession numbers AM748259–AM748262) contains an open reading frame of 6795 bp, encoding a protein consisting of 2265 amino acid residues preceded by a signal peptide of 18 residues (13). The mature secreted protein has a calculated M_r of 244,260 (Fig. 1).

Domain Structure—The domain structures of the new chains are very similar to that of the collagen VI α3 chain (Fig. 2). For comparison with the already known collagen VI chains we use the nomenclature introduced by Chu *et al.* (8). The domains at the N terminus of the collagenous domain are designated with N, the domains at the C terminus of the collagenous domain with C. Numbering starts at the collagenous domain. At the N terminus all three mature proteins contain seven VWA domains (N7–N1), followed by a 336-amino acid residue long collagen triple helical domain. Toward the C terminus they have two VWA domains (C1 and C2) that are followed by a unique sequence (C3) that in the new α6 chain also represents the C-terminal end. In mouse the α4 chain carries a short stretch of 17 amino acid residues at the C-terminal end (C4) that resembles a Kunitz domain. Interestingly, when searching the genomic data bases for exons coding for a complete Kunitz domain, such a domain could be identified at this position in ortholog genes of several species. Only in mouse and rat do the sequences contain a premature stop codon, indicating that, except in rodents, a full Kunitz domain is present at the C terminus of the collagen VI α4 chain (Fig. 3A). In the α5 chain the C-terminal end contains a third VWA domain (C4) followed by another unique domain (C5). A major difference between the new chains and the collagen VI α3 chain is the presence of three additional VWA domains at the N-terminal end of the α3 chain. Interestingly, a splice variant of the collagen VI α3 chain (AAC23667) lacks the first, second, and fourth VWA domains and thereby, as the new chains, contains seven N-terminal VWA domains. The overall identity at the amino acid level is highest between the α5 and α6 chains (44.7%) and lowest between the α4 and α5 chains (28.0%). The overall identity of

the three new chains and the α3 chain varies between 25.9 and 26.7%.

Alternative Splicing—In mouse, two different splice variants of the collagen VI α4 mRNA with premature stop codons can be deduced from EST clones. First, the ESTs AU023415 and BG068629 contain a stop codon in an alternative exon following the exon coding for the N4 domain. If translated, this transcript would yield a protein containing only the first four VWA domains. A second splice variant was detected in the three EST clones BX520360, AI427280, and W48310. Here, an alternative splice donor site in exon 35 coding for the C2 domain and an alternative splice acceptor site in exon 37 coding for the unique domain are used. Due to a shift in codon phase, the new exon codes for a different frame and contains a stop codon 101 bp downstream of the alternative splice site. If translated, this transcript would give a protein that lacks nearly one-half of the C2 domain and the unique domain. Interestingly, the alternative splice site contains a non-canonical GC-AG motif.

A RIKEN cDNA clone coding for the collagen VI α6 chain (AK054356) shows alternative splicing in the 5′-UTR, indicating the presence of two different promoters. Interestingly, due to additional alternative splicing of exon 6, a much shorter open reading frame occurs that would generate a protein containing only the first six VWA domains and lacking the seventh VWA domain, the collagenous domain, and the C-terminal non-collagenous domains.

Analysis of the Collagenous Domains—The 336-amino acid residue long collagenous domains have exactly the same size as that in the collagen VI α3 chain (Fig. 3B). The identity between the collagenous domain of the α3 chain and those in the α4, α5, and α6 chains is 53.3, 49.1, and 51.8%, respectively. A cysteine residue that is also present in the collagenous domain of the collagen VI α3 chain and appears to be involved in tetramer formation and stability (19) is conserved in all new chains.

The locations of the two imperfections in the Gly-Xaa-Yaa repeat found in the collagen VI α3 chain are conserved in all new chains, whereas the α5 and α6 chains have additional imperfections. In both these chains a glycine residue in a Gly-Xaa-Yaa repeat close to the C terminus of the collagenous domain is replaced by a leucine or a valine residue, respectively, introducing another imperfection. Interestingly, the position coincides with an imperfection found in the α1 and α2 chains. In addition, an imperfection is present at the center of the collagenous domains of the collagen VI α5 and α6 chains, where one or two glycine residues of Gly-Xaa-Yaa repeats are lacking, respectively.

In contrast to the collagenous domain of the collagen VI α3 chain, which contains five potentially integrin-binding RGD sequences, in each of the new chains only one RGD motif is present. In the collagen VI α4 and α6 chains the motif is found at exactly the same position where an RGD is present also in the

FIGURE 3. **Amino acid sequence alignment of collagen VI Kunitz domains in different species (A), collagenous domains from the different collagen VI chains in mouse (B), C-terminal parts of the unique domains in mouse (C), and VWA domains containing the metal ion-dependent adhesion site motif in mouse (D).** The sequences for the Kunitz domains of rhesus monkey, dog, and rat were deduced from genomic sequences. The sequences were aligned by CLUSTAL X using the default parameters. The residues forming the trypsin interaction site in the original bovine pancreatic trypsin inhibitor (*BPTI*) (34) are marked with a *number sign*, the cysteine residues with *asterisks*, and the RGD sequences with *dots*. Imperfections in the collagenous domains are *boxed* and numbered *I1–I4*. The conserved metal ion-dependent adhesion site (42) and the conserved hydrophobic moieties (43) are denoted with ▼ and φ, respectively.

collagen VI $\alpha3$ chain (Fig. 3*B*). The content of proline or hydroxyproline in the *X* and *Y* positions is lower (17.4 –20.5%) than in the fibril-forming collagen I $\alpha1$ or collagen II $\alpha1$ chains

(26). N- and C-terminal of the collagenous domains several cysteine residues are present, which might form intermolecular disulfide bridges that enhance the stability of the trimeric col-
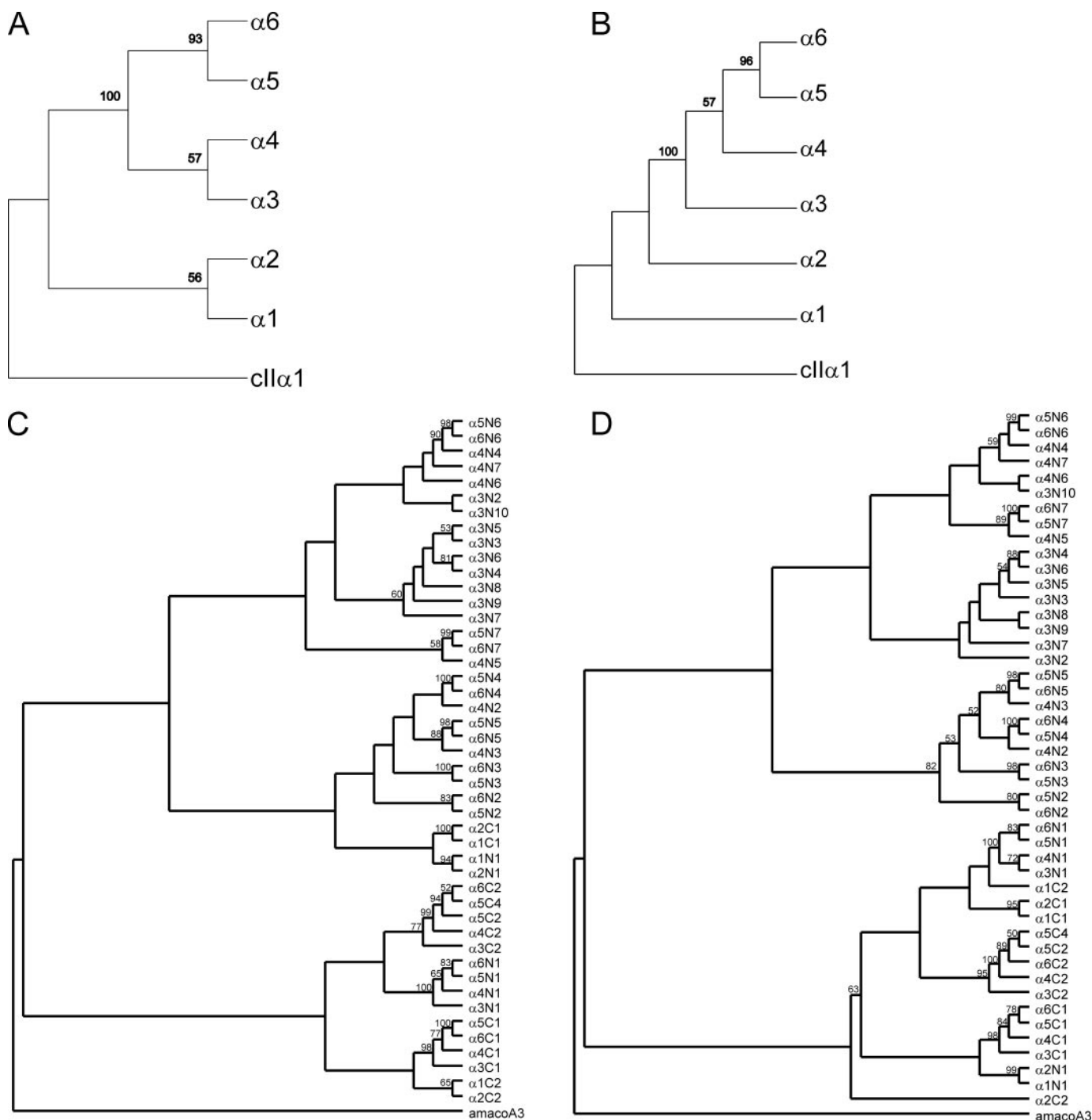
FIGURE 4. **Phylogenetic trees of the collagenous domains (*A* and *B*) and the VWA domains (*C* and *D*) of collagen VI.** The sequences from all six collagen VI chains in mouse were aligned using CLUSTAL X. The trees were constructed using the programs PROTEIN DISTANCE, NEAREST NEIGHBOUR, and CONSENSE (*A* and *C*) and PROTEIN PARSIMONY and CONSENSE (*B* and *D*) of the PHYLIP package version 3.66. Bootstrap analyses using 100 replicates were performed to show the significance. The *numbers* indicate the statistical weight of the individual branches. The collagenous domain of the collagen II α1 chain (*cIIα1*) (*A* and *B*) and the VWA3 domain of AMACO (44) (*C* and *D*) were used as outgroups.

lagens. In phylogenetic analyses using protein distance and protein parsimony, the collagenous domains of the α3, α4, α5, and α6 chains group in one clade (Fig. 4, *A* and *B*).

Analysis of the VWA Domains—Of the 28 VWA domains present in the new collagen VI chains, the metal ion-dependent adhesion site (D*XSXSX*nT*X*nD, where n represents a variable number of amino acid residues) motif, is fully conserved only in 8 (Fig. 3*D*). Sequence alignment of the VWA domains of the new

chains with their counterparts present in the collagen VI α1–α3 chains highlights the homology (Fig. 3*D* and supplemental Fig. 1). The highest sequence identity between two VWA domains of the new chains is 92.1% for α5N7 and α6N7. High identity values were also obtained for the α5N4 and α6N4 (64.5%), α5N5 and α6N5 (51.9%), α5C2 and α6C3 (52.9%), α5C1 and α6C1 (50.5%), and α5N1 and α6N1 (50.3%). Among the various VWA domains found in the collagen VI α1–α3 chains, the N10
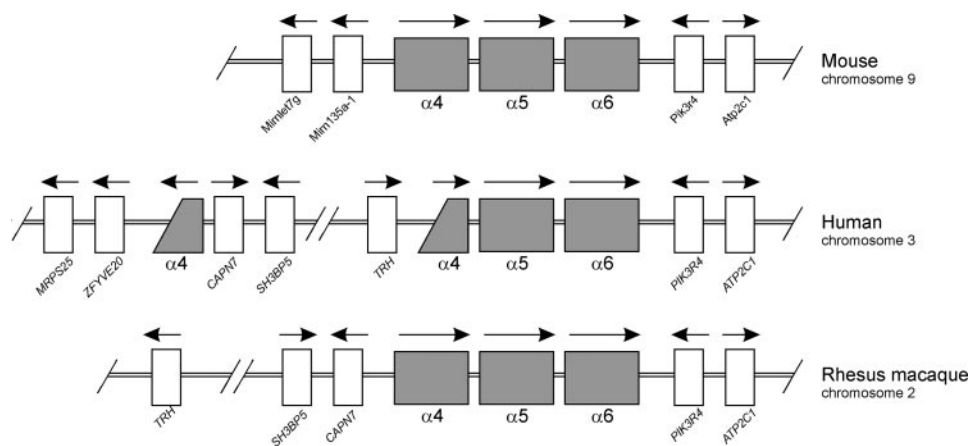
FIGURE 5. **Loci for collagen VI α4, α5, and α6 in the genomes of man, rhesus monkey, and mouse.** The orientation of the genes is indicated by *arrows*.

(α5C3 22/32 (68%)), protein-tyrosine phosphatase (α5C5, 34/71 (47%)), and dynein cytoplasmic 2 heavy chain 1 (α6 26/60 (43%)).

Structure of the Murine Collagen Col6a4-Col6a6 Genes—The new mouse collagen VI genes map to chromosome 9 (9F1) (Fig. 5). The genomic sequences are completely contained in the public data bases (NT_039477 and NW_001030918). The genes lay head to tail in tandem orientation on the minus strand. The *Pik3r4* gene and the *Mirn135a1* gene are located downstream and upstream of the new collagen genes, respectively. We identified exons by flanking consensus splice signals and by comparison with the respective cDNAs. The exon/intron organization of the three genes is very similar (Fig. 6 and supplemental Tables 2–4) regarding size, exon and intron length, and codon phase. The *Col6a4* and *Col6a5* genes are 112 kb, and the *Col6a6* gene is 104 kb long. They consist of 38, 44, and 37 exons, respectively, that code for the translated part of the mRNA (Fig. 6). The first exon in each gene completely encodes the 5′-UTR. All second exons code for the signal peptide sequence followed by six exons coding for the first six VWA domains (N7–N2), whereas the VWA domain N1 is encoded by three exons. The collagenous domains are encoded by exons 12–30. Interestingly, intron 24 of the *Col6a4* gene is a GC-AG-type intron. Exons 31 and 32 code for short spacer regions. The VWA domains C1 and C2 are encoded by exons 33/34 and 35, respectively. The structures of the three genes differ at the 3′-end. In *Col6a4* the unique sequence is encoded by two exons followed by a last exon coding for the truncated Kunitz domain and the 3′-UTR. In *Col6a6* the last two exons code for the unique domain and 3′-UTR. The more complex structure of the C-terminal end of the collagen VI α5 chain is also reflected in the *Col6a5* gene structure. The unique domain between the VWA domains C2 and C4 is encoded by two exons followed by the exon coding for the additional VWA domain C4. As in *Col6a6*, the last two exons of *Col6a5* code for the unique domain and 3′-UTR. Although there is only partial homology between the unique domains of the new collagen chains, each unique domain is encoded by two exons where the first exon is always about 95 bp and the second about 200 bp long, pointing to the likelihood of a common ancestor.

domain of the collagen VI α3 chain shows the highest identity value to the N7 domain of the α4 chain (39.5%). Similar identity values were obtained for α3N9 and α4N7 (34.7%) and α3C1 and α4C1 (34.5%). Identity values between the α3 chain VWA domains and α5 and α6 chain VWA domains are not higher than 28.4 and 28.9%, respectively. The identity between the VWA domains of the new chains and those of the collagen VI α1 and α2 chains is always lower than 24.0%. In phylogenetic analyses using protein distance and protein parsimony, all the VWA domains of the α5 and α6 chains pair up together (Fig. 4, *C* and *D*). The C-terminal VWA domains of the α3, α4, α5, and α6 chains group to a distinct branch in which the C1 domains are in one subbranch and the C2 domains and the C4 domain of the α5 chain are in another. Similarly, the N1 domains of the α3, α4, α5, and α6 chains all cluster together (Fig. 4, *C* and *D*).

Analysis of the Unique Domains—The unique sequences at the C-terminal end follow directly after the second C-terminal VWA domains (C2). In the collagen VI α5 chain a second unique domain is present C-terminal of the C4 domain. The unique domains are 99–111 amino acid residues long. The unique domain of the α4 chain and the first unique domain of the α5 chain as well as the second unique domain of the α5 chain and the unique domain of the α6 chain share some pairwise similarity, 31.6 and 26.1%, respectively. However, a stretch of 15 amino acid residues at the beginning of each domain is highly identical in all four unique domains and has a cysteine residue at the end (Fig. 3*C*). Interestingly, the unique sequence of the collagen VI α3 chain, C-terminal to the C2 domain, also shares some homology to the unique domains of the new chains, most clearly in the C-terminal portions, and particularly the cysteine residue is conserved (Fig. 3*C*). Interestingly, shortly after the highly homologous stretch, an RGD motif is present in both the α4 chain and the first unique domain of the α5 chain, whereas this motif is missing in the α6 chain and in the second unique domain of the α5 chain (Fig. 3*C*). In addition to the single RGD motifs present in each of the collagenous domains, these two RGD motifs are the only ones found in the new collagen VI chains. An RGD motif is lacking in the unique domain of the collagen VI α3 chain. BLAST searches with the unique sequences revealed some weak homologies to intracellular proteins like the REST corepressor 1 (α4 35/83 (42%)), ubiquitin D

New Collagen VI Genes in Man—The orthologs of the new mouse genes map to human chromosome 3q21 (Fig. 5). The tandem orientation of the genes is conserved, but the gene coding for the α4 chain is broken into two pieces, and the 5′ region of the gene is located at 3p24.3. Only the region downstream of the new collagen VI genes, coding for *PIK3R4*, is in synteny in man and mouse. The breakpoint resembles the large scale pericentric inversion that occurred in the common ancestor of the African apes and is present in modern human chromosome 3 as well as in the chimpanzee and gorilla orthologs, but not in orangutan or Old World monkeys (27). In contrast to rhesus
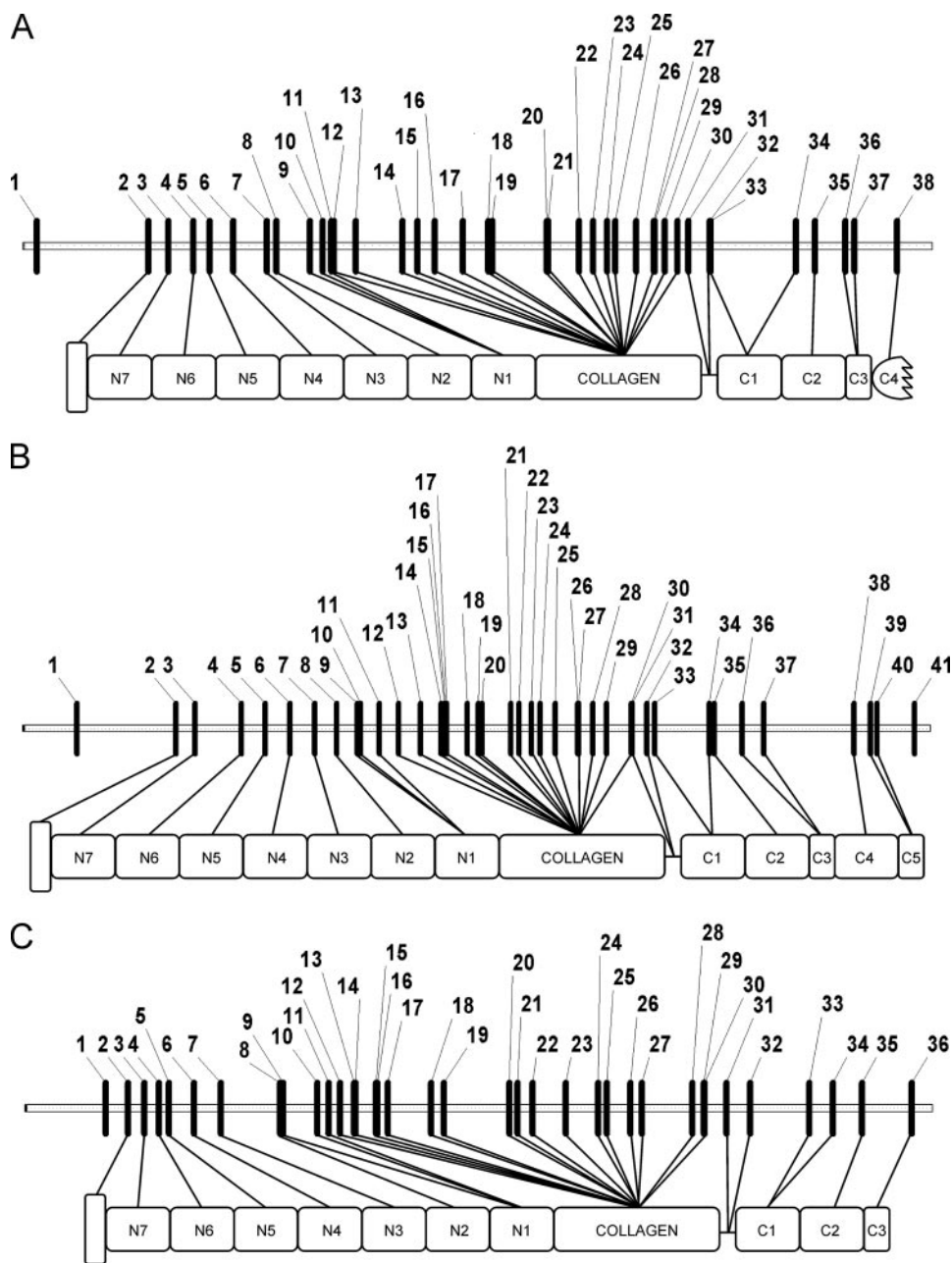
FIGURE 6. **Organization of the murine *Col6a4*, *Col6a5*, and *Col6a6* genes.** The *numbers* indicate the positions of exons. The protein domain structure is given below, and the corresponding exons are indicated by *lines*.

sequence. A 32-amino acid long proline-rich stretch at the C terminus of C1 is missing in man, and the unique domains are highly divergent. In addition, at two positions in the C-terminal part an amino acid residue is deleted and at three positions an amino acid residue is inserted into the human α5 chain. Most of the cysteine residues are conserved, but there is an additional cysteine present in the collagenous domain of the human α5 chain. However, the cysteine codon resembles an SNP (rs1497312) leading to a non-synonymous exchange to a serine codon. The positions and sizes of the imperfections in the collagenous domain are identical to those in mouse, whereas the RGD motif in the collagenous domain of the α5 chain is lost. Instead there is a new RGD motif at the N terminus of the collagenous domain. The two RGD motifs present in the unique domain of mouse are also missing in man. Another SNP (rs11355796), which resembles the deletion of a thymidine at the C terminus forms a premature stop codon, leading to a full-length protein of 2590 residues (supplemental Fig. 2). No information is available on the population frequency, but both variants are found in the TRACES-WGS data base. Interestingly, in the alternative Celera assembly of the human genome the deletion is present, whereas the thymidine is found in the reference assembly, leading to a longer protein. In contrast to murine *Col6a5*, human *COL6A5* contains an additional intron in the 3'-UTR that leads to three different C termini by alternative splicing, thus resulting in full-length α5 chain isoforms of 2526, 2614, or 2615 amino acid residues (supplemental Fig. 2). The human collagen VI α6 chain has an identity of 83.4% at the amino acid level to the mouse ortholog, and only the last 30 amino acids at the C terminus show some differences (supplemental Fig. 3). The positions of the signal peptide cleavage site, the RGD motif, and all cysteine residues in the mature protein are completely conserved.

macaque, the human *COL6A4* is interrupted after the first exon coding for the collagenous domain, and EST clones for both parts of the gene can be found in the data bases. However, due to the presence of stop codons that are distributed over the sequence, both parts of the human *COL6A4* are likely to be transcribed non-processed pseudogenes. The corresponding cDNAs of the human *COL6A5* and *COL6A6* were cloned as overlapping partial clones by RT-PCR using primers deduced from the genomic sequence and sequenced (accession numbers AM774225–AM774227 and AM906078–AM906084). The human collagen VI α5 chain has an identity of 73.1% at the amino acid level to the mouse ortholog (supplemental Fig. 2). The non-identities are not evenly distributed over the

Interestingly, variants of the human *COL6A5* gene were recently shown to be associated with atopic dermatitis (28). The authors of that study designated the human *COL6A5* as *COL29A1*. Although not present in the original publication, the sequence was recently published in the data base (accession

number EU085556). Although in the publication the length of the protein sequence was given as 2614 amino acids, the sequence submitted to the data base is 2615 amino acids long. The reason for this difference is unclear, but the amino acid sequence is nearly completely identical with the third variant presented here (supplemental Fig. 2), the only difference being that residue 2560 is a serine residue instead of the asparagine found by us. Nevertheless, the second splice variant presented here contains 2614 amino acids. Surprisingly, the 5′-UTR region of the EU085556 contains a duplication of the 19-bp sequence GTGCGGCGCGGACCAGGGC that is not present in our sequence and is found neither in the alternative Celera assembly of the human genome nor in any of the 21 TRACE-WGS clones that cover this region but is present in the reference assembly of the human genome.

Expression of the New Mouse Collagen VI Genes—To determine the length of the new collagen VI mRNAs we performed Northern hybridization with total RNA or mRNA (Fig. 7A). The mRNA coding for the α6 chain was readily detected as a 9.7-kb band in total RNA derived from the lungs of newborn mice. Several messages coding for the α4 chain were detected in total RNA derived from the same source, probably indicating alternative splicing. The most prominent mRNA band had a length of 8.4 kb whereas weaker bands appeared at 11.7, 6.7, and 5.0 kb. A 9.5-kb message coding for the α5 chain was detected in purified mRNA derived from sterna of 4-week-old mice. RT-PCR was performed to screen the tissue distribution of the new collagen VI chains (Fig. 7, B and C). Products corresponding to the mRNAs for the α5 and α6 chains could be detected in lung, heart, kidney, muscle, brain, intestine, skin, femur, and sternum of newborn mice. In addition, α6 chain mRNA could be detected in calvaria. The α4 chain mRNA shows a more restricted tissue distribution and could be detected in lung, kidney, brain, intestine, skin, sternum, and weakly in calvaria (Fig. 7B). In adult mice, expression of the α4 chain is lost in most tissues, and RT-PCR showed a signal only in ovary and very weakly in spleen, lung, uterus, and brain. In contrast, the α5 chain is widely expressed also in adult mice, and mRNA could be detected in lung, heart, kidney, spleen, muscle, ovary, uterus, brain, skin, liver, and sternum, whereas the α6 chain expression is more restricted and could be detected in lung, heart, muscle, ovary, brain, liver, and sternum (Fig. 7C).

The New Collagen VI Chains Copurify with α1, α2, and α3 Chain-containing Collagen VI Prepared from Newborn Mice—If the new collagen VI chains assemble with known collagen VI chains, they should be present in conventional collagen VI preparations. Thus we isolated collagen VI from newborn mouse carcasses (25) and tested this preparation for the presence of the new chains by immunoblot. For this purpose, we generated antibodies specific for the new chains. Tagged versions of different N-terminal VWA domains were recombinantly expressed in EBNA293 cells, and the recombinant proteins were purified by affinity chromatography and used to immunize rabbits. The antisera were affinity-purified before use, and cross-reactivity among the new collagen VI chains was tested by ELISA (Fig. 8 and not shown). All three new chains were detected after reduction of the collagen VI preparation as major bands running above the 220-kDa marker (Fig. 8), con-
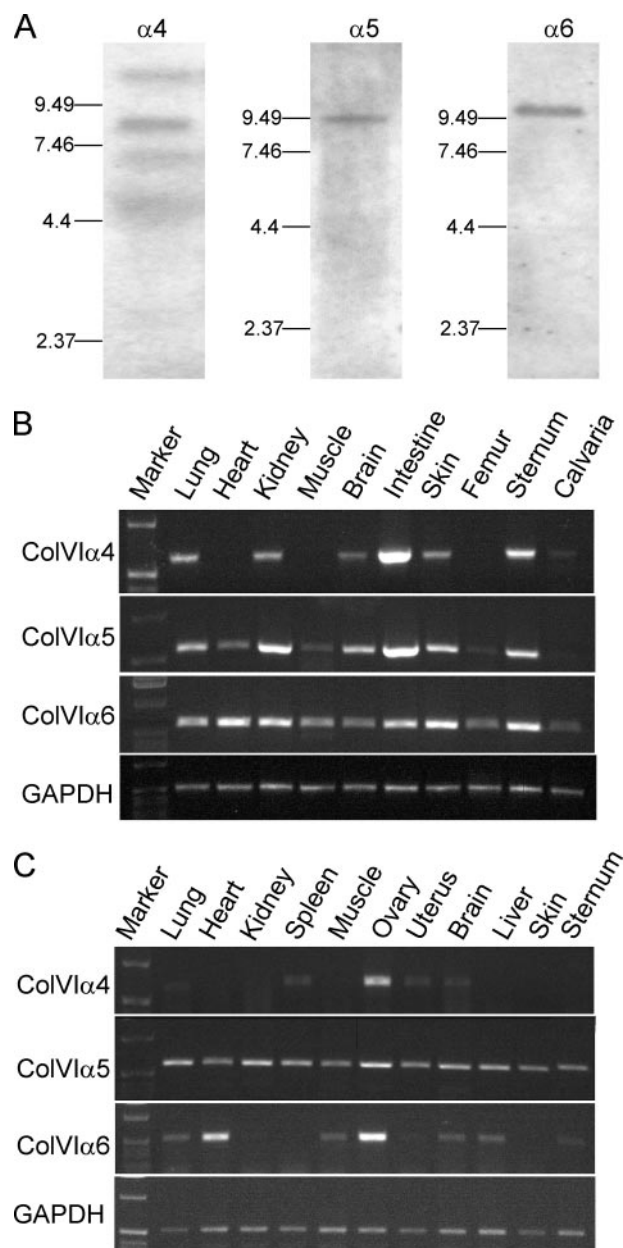

FIGURE 7. **Northern blot (A) and RT-PCR (B and C) analysis of the new mouse collagen VI chain mRNA species.** A, Northern hybridization was performed for the collagen VI α4 and α6 chains with 5 μg of total RNA from lung of newborn mice and for the α5 chain with of 1 μg of poly(A)+ RNA from sternum of 4-week-old mice. Probes were generated using primers α4m8 and α4m9 for the α4 chain, α5m2 and α5m7 for the α5 chain, and α6m6 and α6m10 for the α6 chain. Position of size markers are indicated on the *left*. B and C, RT-PCR analysis was performed using primer pairs α4m6 and α4m7 for the α4 chain, α5m4 and α5m5 for the α5 chain, and α6m2 and α6m9 for the α6 chain. Template RNA was isolated from newborn (B) and adult mice (C). The 1-kb ladder from Invitrogen was used as a reference.

sistent with the calculated molecular masses. For the collagen VI α4 and α5 chains additional lower migrating bands were detected (Fig. 8), indicating alternative splicing or proteolytic processing. The weak smear with lower mobility seen for the α4 chain (Fig. 8) could indicate the presence of non-reducible cross-linked molecules.

The Collagen VI α5 and α6 Chains, but Not the α4 Chain, Are Deposited in the Extracellular Matrix of Skeletal Muscle—Mutations in collagen VI lead to muscular dystrophies in humans,
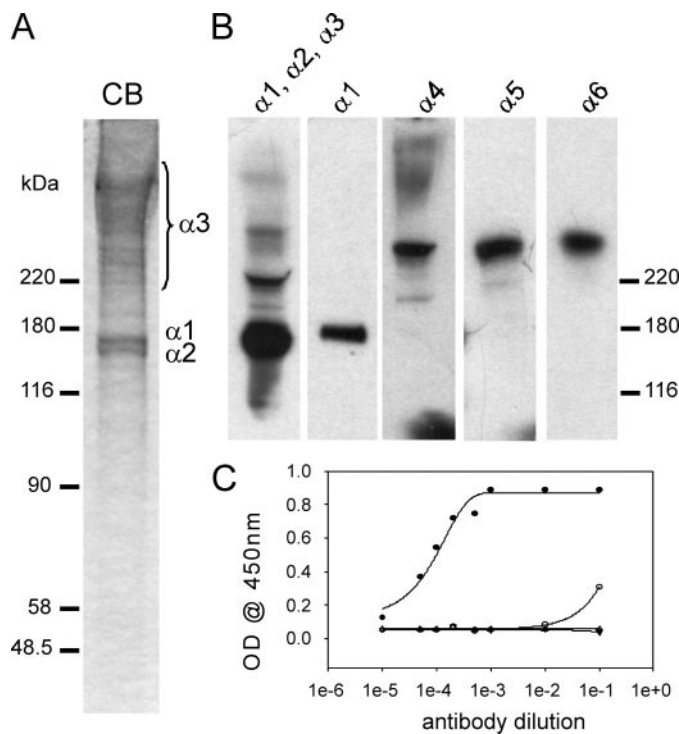
FIGURE 8. **Analysis of collagen VI purified from newborn mice (*A* and *B*) and ELISA of antibody reactivity against the collagen VI α6 chain (*C*).** Carcasses of newborn mice were extracted, and collagen VI was isolated by molecular sieve column chromatography as described previously (25). The proteins were submitted to SDS-PAGE on 4–12% polyacrylamide gradient gels under reducing conditions. *A*, Coomassie Brilliant Blue-stained gel. *B*, immunoblot of the same preparation with antibodies against collagen VI (α1, α2, α3; *lane 1*), the collagen VI α1 chain (*lane 2*), and the new collagen VI chains α4, α5, and α6 (*lanes 3–5*). *C*, reactivity of affinity-purified antibodies against the collagen VI α6 chain and lack of reactivity of antibodies against the α4 and α5 chains against the collagen VI α6 chain N2–N6 domains. ELISA plates were coated with 0.5 μg of recombinant collagen VI α6 N2–N6 protein per well, and antibodies were diluted as indicated. *Filled circles*, rabbit α6 antibody; *open circles*, rabbit α5 antibody; *filled triangles*, rabbit α4 antibody; *open triangles*, guinea pig α4 antibody. *OD*, optical density.

and mice lacking collagen VI display a myopathic phenotype affecting skeletal muscle (20). We therefore tested cryostat sections of adult mouse quadriceps femoris muscle for the expression of the new collagen VI chains (Fig. 9). By immunohistochemistry using the polyclonal antibodies, collagen VI α5 and α6 were readily detected in skeletal muscle of adult mice. As the specific antibodies against the collagen VI α4 chain did not stain skeletal muscle we tested its reactivity on sections of small intestine where this chain could be strongly detected below the mucosal layer (Fig. 9). The strong reactivity of the antibody with intestine indicates that the collagen VI α4 chain is truly absent from skeletal muscle.

The targeted interruption of the *Col6a1* gene in mouse completely abolishes the secretion of the collagen α2 and α3 chains (20), showing the need for a heterotrimeric assembly to form a functional collagen VI molecule. To determine the influence of the lack of the collagen VI α1 chain on the assembly of the new collagen VI chains we analyzed their occurrence in α1 chain-deficient mice (20). The new collagen VI chains could not be detected in quadriceps femoris of collagen VI α1 chain-deficient mice by immunohistochemistry (Fig. 9), indicating a participation of the α1 chain in the assembly of collagen VI mole-

cules containing the new chains. The absence of the α5 and α6 chains in *Col6a1* knock-out mice was also confirmed by immunoblot analysis of diaphragm extracts (Fig. 10). For wild type mice, incubation with antibodies specific for either the collagen VI α5 or α6 chain resulted in clearly identifiable bands above 220 kDa. When the same method was applied to diaphragm from collagen VI α1 chain-deficient mice, no bands were detected, supporting the results from the immunohistochemical analysis.

DISCUSSION

We report on the identification and initial biochemical characterization of three new collagen VI chains, named collagen VI α4, α5 and α6. The mouse *Col6a4–Col6a6* genes are arranged in tandem on chromosome 9 and were numbered according to their appearance from 5′ to 3′ on the coding strand. Cloning of the cDNAs by RT-PCR and immunohistochemistry and immunoblot using antibodies raised against recombinant fragments confirmed the expression of the new collagen VI genes in mouse. These genes have previously only been incompletely annotated or incorrectly predicted by conceptual translation or gene prediction programs.

The sequences and the domain structures of the new proteins show that they represent new collagen VI chains, which probably occur as the consequence of a gene duplication of the common ancestor of the collagen VI α3 and of the new α chain genes, followed by additional duplications. The identical size of the collagenous domains of the new α chains, compared with that of the α3 chain, implies that they could substitute for the α3 chain, probably forming α1α2α4, α1α2α5, or α1α2α6 heterotrimers. The close relation between the α3 chain and the new chains is also reflected by the almost identical exon/intron organization of the portions of the respective genes encoding the collagenous domains. With the exception of the last exon, the exon sizes are the same. Alternative splicing has been reported for the α3 chain mRNA (29) leading to production of shorter protein isoforms with molecular sizes similar to those of the new chains. The collagen VI arrangement known to date is composed of α1, α2, and α3 chains that associate intracellularly in a stoichiometric ratio to form triple helical monomers. Monomers then assemble into dimers and tetramers, which are finally secreted and deposited in the extracellular matrix where they form beaded filaments by interactions of their non-collagenous domains (3). There is good evidence that the α3 chain expression is essential for the formation of functional collagen VI molecules, as human SaOS-2 cells that are deficient in α3 chain expression do not produce triple helical collagen VI (30). Although the length of the collagenous domain of the collagen VI α1 chain is also identical to that of the new chains and the α2 chain is only one amino acid residue shorter, there are other criteria that clearly show the closest relationship of the new chains with the α3 chain. First, the exact position of the cysteine residue within the collagenous domain is conserved in the α3, α4, α5, and α6 chains. In the α3 chain this cysteine appears to be involved in tetramer formation and stability (19). The α1 and α2 chains also contain a cysteine each, but these are at a different position, and they appear to be involved in the stabilization of the supercoil that is formed during antiparallel dimer forma-
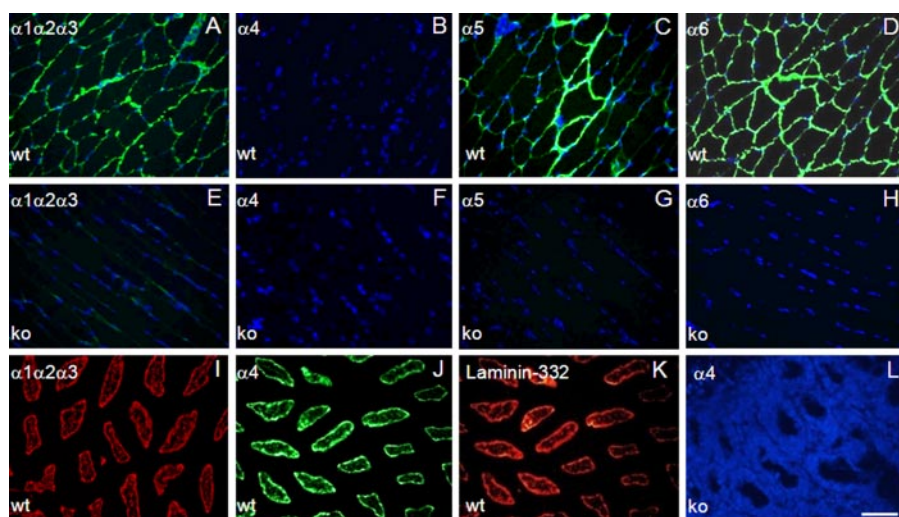
FIGURE 9. **Immunohistochemical analysis of wild type and *Col6a1* knock-out mouse.** Immunohistochemistry was performed on frozen sections from mouse quadriceps femoris muscle (*A–H*) and small intestine (*I–L*) from wild type (*wt*) (*A–D* and *I–K*) or *Col6a1* knock-out mice (*ko*) (*E–H* and *L*). Sections were incubated with the affinity-purified antibodies against the collagen VI α4 N3–N6 domain (*B*, *F*, *J*, and *L*), the collagen VI α5 N3 domain (*C* and *G*), the collagen VI α6 N2–N6 domain (*D* and *H*), human collagen VI from placenta (detecting the classical collagen VI chains α1α2α3; *A*, *E*, and *I*) and laminin 332 (*K*), followed by an AlexaFluor 488-conjugated goat anti-rabbit IgG (*green*, *A–H*), AlexaFluor 546-conjugated goat anti-rabbit IgG (*red*, *I* and *K*), or AlexaFluor 488-conjugated goat anti-guinea pig IgG (*green*, *J* and *L*). Antibodies against the classical collagen VI chains (α1α2α3) and the α5 and α6 chains, but not such against the α4 chain, strongly stained the extracellular matrix surrounding the muscle fibers of wild type (*A–D*) mice. In small intestine, antibodies against classical collagen VI (α1α2α3) (*I*) and antibodies against the collagen VI α4 chain show co-staining with those against the basement membrane marker laminin-332 (*K*). In collagen VI α1 chain-deficient mice, staining for the new collagen VI chains is absent (*L*). Nuclei were counterstained with 4′,6-diamidino-2-phenylindole (*blue*, *A–H* and *L*). The bar is 100 μm.

tion (31). Second, it has also been suggested that the supercoiled dimer is partially stabilized by ion pairs between different segments along the supercoil (32). In the α1α2α3 triple helical monomer the supercoiled part of the α1 chain carries a high negative net charge, whereas that of the α3 chain has a high positive net charge and that of the α2 is close to neutral. All three new chains carry a positive net charge that is even higher than that of the α3 chain. In addition, the positions of the two Gly-Xaa-Yaa imperfections present in the α3 chain, giving the supercoil a clearly segmented character (32), are conserved. Third, phylogenetic analyses based on the collagenous domains, using protein distance and protein parsimony methods, clusters the new chains with the collagen VI α3 chain, whereas the α1 and the α2 chains form a different branch (Fig. 4, *A* and *B*). In addition, the VWA domains C1 and C2 of the α3, α4, α5, and α6 chains also cluster in common branches (Fig. 4, *C* and *D*).

All three new chains could be detected in collagen VI preparations from newborn mouse carcasses. The Coomassie-stained gel clearly showed the distinct α1 and α2 bands and a very heterogenous distribution of bands above 220 kDa (Fig. 8*A*). By immunoblot using affinity-purified antibodies, the new chains were identified as a part of this purified collagen VI preparation (Fig. 8*B*), indicating that trimeric assemblies containing the new chains are present in tissues. Further experimental evidence for an assembly of the new chains into collagen VI came from the study of their expression in *Col6a1* knock-out mice. It was shown earlier that the absence of the collagen VI α1 chain also leads to the lack of the secretion of the α2 and α3 chains (20), indicating that the α1 chain is essential for the assembly of

collagen VI molecules. The complete lack of the new chains in *Col6a1* knock-out mice clearly shows that the presence of the collagen VI α1 chain is a prerequisite also for their secretion. This observation strongly suggests that the new chains assemble in a similar way as proposed for the α1, α2, and α3 chain-containing collagen VI.

It has been shown that the N5 and the C5 domains of the collagen VI α3 chain are critical for the microfibril formation (9, 33). Based on the sequence information alone it is not clear which domains in the new chains correspond to the N5 domain of the α3 chain. In some species only the α4 chain contains a domain distinctly homologous to the α3 chain C5 domain. As in the α3 chain, this domain resembles a Kunitz domain. Interestingly, the Kunitz domain of the α4 chain is truncated in mouse and rat, but the N-terminal part, which contains the trypsin interaction site in the original bovine pancreatic trypsin inhib-

itor (34), is still present in the truncated molecules and could serve as an interaction module. The α5 and α6 chains lack a Kunitz domain, which may indicate differences in the assembly of α5 and α6 chain-containing microfibrils. Indeed, it will be interesting to study how collagen VI of different composition assembles. Do fibrils contain α1, α2, and only one of the four related α3, α4, α5, and α6 chains or are mixed assemblies possible? The latter alternative would lead to a very high number of possible permutations.

The ortholog human genes are present on a, in an evolutionary context, very interesting part of chromosome 3. A large pericentric inversion occurred some time after the split of Homininae and Ponginae (27). The 3′ breakpoint of the inversion is located within *COL6A4* and leads to its inactivation. Although both parts of *COL6A4* are still present and can be easily identified by their sequence, both have become transcribed non-processed pseudogenes. Thereby Homininae have become natural *COL6A4* knock-outs (Fig. 5). This raises the question of whether one of the remaining genes has taken over the function of the lost one. The major structural difference between the collagen VI α4 chain and the α5 and α6 chains is at the C terminus, where the fibronectin type III domain and the Kunitz domain occur only in the α4 chain. However, when comparing human and mouse collagen VI α5 and α6 chains, higher divergence is found at the C terminus of the α5 chain, which in addition shows alternative splicing and could represent an adaptation to a need to replace the α4 chain in Homininae.

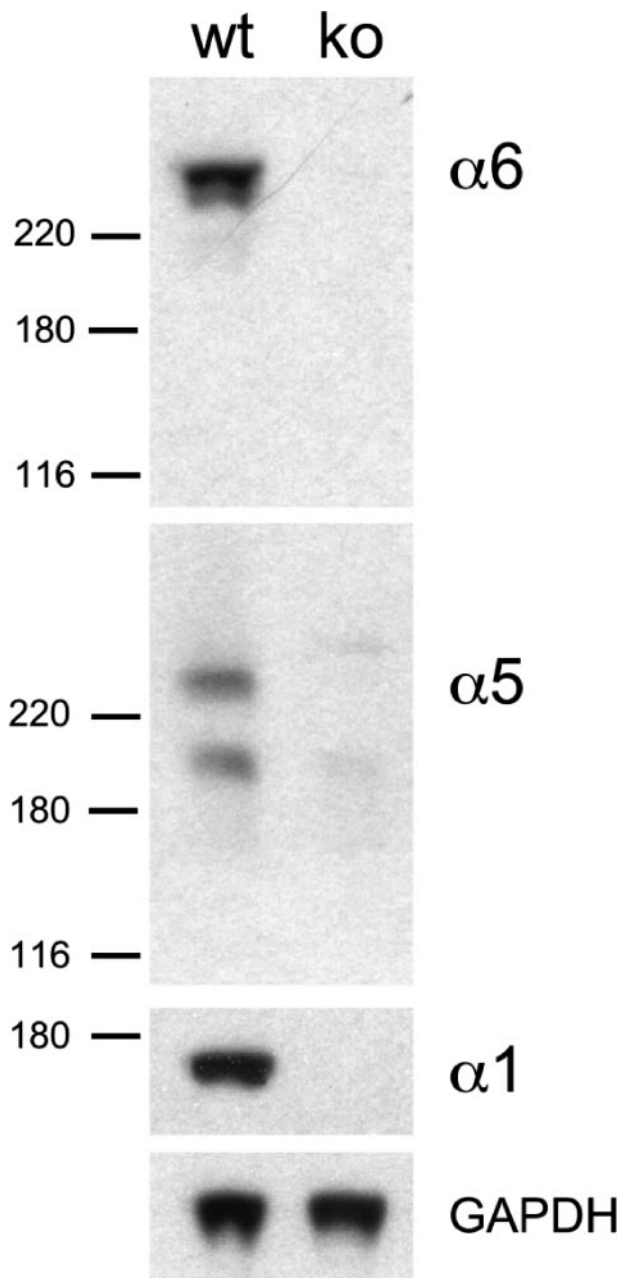Recently, the human *COL6A5* was associated with atopic dermatitis in a linkage study and designated *COL29A1* (28).

FIGURE 10. **Immunoblot analysis of collagen VI $\alpha5$ and $\alpha6$ chains in wild type and *Col6a1* knock-out mouse diaphragm extracts.** Diaphragm muscle of adult wild type (*wt*) and *Col6a1* knock-out mice (*ko*) were extracted with lysis buffer (see "Materials and Methods"). The proteins were submitted to SDS-PAGE with prior reduction on 4–12% polyacrylamide gradient gel, transferred to a membrane, and immunostained with specific antibodies against the collagen VI $\alpha1$, $\alpha5$, and $\alpha6$ chains or glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*).

Although neither DNA nor protein sequences were originally published, later the cDNA sequence became accessible in the data base. The protein sequence is, except for a single amino acid exchange, identical to the third potential splice variant presented here. As we have shown for mouse by RT-PCR, the collagen VI $\alpha5$ chain is expressed in skin in man also (28). The expression in other tissues only partially overlaps; however intestine and lung also show expression in both species. A variety of nonsynonymous coding SNPs were described, but none could explain the association of *COL6A5* with atopic dermatitis

on its own. It was therefore proposed that several variants or combinations associated with the most common haplotype of *COL6A5* are involved in the etiology of the disease. In addition, a strongly maternal transmission pattern was found, which could be due either to imprinting or to maternal effects through an interaction of the child's genotype with the maternal environment during prenatal life. Another susceptibility locus for atopic dermatitis is linked to 3p24-22 (35), which is exactly the breakpoint area where the 5′ part of the *COL6A4* pseudogene is located. It could be that the mechanism leading to atopic dermatitis is more complex and that the expression of the non-processed $\alpha4$ chain pseudogenes by a yet unknown mechanism influences $\alpha5$ chain expression. The maternal transmission pattern could point to such a mechanism. A number of pseudogenes have been described where gene conversion between a functional copy of a gene and a neighboring pseudogene causes disease (36). However, in the present case the mechanism is likely to be more complex.

The newly identified locus for atopic dermatitis in *COL6A5* could correlate to another susceptibility locus found on chromosome 21p21 in a Swedish population that may contain a susceptibility gene modulating the severity of atopic dermatitis especially in combination with asthma (37). Both 21p21 and 3p24 have also been described as asthma susceptibility loci (38). Interestingly, *COL6A1* and *COL6A2* are located on 21p21, which could point to a more general role of collagen VI in the development of atopic dermatitis or asthma. In contrast, there is clear evidence for the role of collagen VI in the etiology of Bethlem myopathy and Ullrich congenital muscular dystrophy. A variety of mutations in all so far known collagen VI chains have been identified (for review see Ref. 19). Interestingly, patients who have phenotypes typical of Bethlem myopathy or Ullrich congenital muscular dystrophy, but in whom mutations in the collagen VI $\alpha1$, $\alpha2$ and $\alpha3$ chains could not be detected have been described (39–41). It is therefore tempting to speculate that mutations in the new collagen VI chains may cause muscular disease. Indeed, in mouse skeletal muscle the affinity-purified antibodies detected two of the new chains, $\alpha5$ and $\alpha6$, but not $\alpha4$, associated with the extracellular matrix. In contrast, the $\alpha4$ chain was absent from mouse muscle but detected close to the basement membrane underlying the mucosal cell layer in small intestine. This differential distribution indicates that the new chains may have tissue-specific functions allowing a modulation of collagen VI properties.

REFERENCES

1. Myllyharju, J., and Kivirikko, K. I. (2004) *Trends Genet.* **20,** 33–43
2. Veit, G., Kobbe, B., Keene, D. R., Paulsson, M., Koch, M., and Wagener, R. (2006) *J. Biol. Chem.* **281,** 3494–3504
3. Furthmayr, H., Wiedemann, H., Timpl, R., Odermatt, E., and Engel, J. (1983) *Biochem. J.* **211,** 303–311
4. Colombatti, A., Bonaldo, P., Ainger, K., Bressan, G. M., and Volpin, D. (1987) *J. Biol. Chem.* **262,** 14454–14460
5. Engvall, E., Hessle, H., and Klier, G. (1986) *J. Cell Biol.* **102,** 703–710
6. Whittaker, C. A., and Hynes, R. O. (2002) *Mol. Biol. Cell* **13,** 3369–3387
7. Chu, M. L., Pan, T. C., Conway, D., Kuo, H. J., Glanville, R. W., Timpl, R., Mann, K., and Deutzmann, R. (1989) *EMBO J.* **8,** 1939–1946
8. Chu, M. L., Zhang, R. Z., Pan, T. C., Stokes, D., Conway, D., Kuo, H. J., Glanville, R., Mayer, U., Mann, K., Deutzmann, R., and Timpl, R. (1990) *EMBO J.* **9,** 385–393

9. Fitzgerald, J., Morgelin, M., Selan, C., Wiberg, C., Keene, D. R., Lamande, S. R., and Bateman, J. F. (2001) *J. Biol. Chem.* **276,** 187–193
10. Ball, S. G., Baldock, C., Kielty, C. M., and Shuttleworth, C. A. (2001) *J. Biol. Chem.* **276,** 7422–7430
11. Ball, S., Bella, J., Kielty, C., and Shuttleworth, A. (2003) *J. Biol. Chem.* **278,** 15326–15332
12. Sipila, L., Ruotsalainen, H., Sormunen, R., Baker, N. L., Lamande, S. R., Vapola, M., Wang, C., Sado, Y., Aszodi, A., and Myllyla, R. (2007) *J. Biol. Chem.* **282,** 33381–33388
13. Bonaldo, P., Russo, V., Bucciotti, F., Doliana, R., and Colombatti, A. (1990) *Biochemistry* **29,** 1245–1254
14. Bidanset, D. J., Guidry, C., Rosenberg, L. C., Choi, H. U., Timpl, R., and Hook, M. (1992) *J. Biol. Chem.* **267,** 5250–5256
15. Brown, J. C., Golbik, R., Mann, K., and Timpl, R. (1994) *Matrix Biol.* **14,** 287–295
16. Tillet, E., Wiedemann, H., Golbik, R., Pan, T. C., Zhang, R. Z., Mann, K., Chu, M. L., and Timpl, R. (1994) *Eur. J. Biochem.* **221,** 177–185
17. Finnis, M. L., and Gibson, M. A. (1997) *J. Biol. Chem.* **272,** 22817–22823
18. Wiberg, C., Klatt, A. R., Wagener, R., Paulsson, M., Bateman, J. F., Heinegard, D., and Morgelin, M. (2003) *J. Biol. Chem.* **278,** 37698–37704
19. Lampe, A. K., and Bushby, K. M. (2005) *J. Med. Genet.* **42,** 673–685
20. Bonaldo, P., Braghetta, P., Zanetti, M., Piccolo, S., Volpin, D., and Bressan, G. M. (1998) *Hum. Mol. Genet.* **7,** 2135–2140
21. Irwin, W. A., Bergamin, N., Sabatelli, P., Reggiani, C., Megighian, A., Merlini, L., Braghetta, P., Columbaro, M., Volpin, D., Bressan, G. M., Bernardi, P., and Bonaldo, P. (2003) *Nat. Genet.* **35,** 367–371
22. Angelin, A., Tiepolo, T., Sabatelli, P., Grumati, P., Bergamin, N., Golfieri, C., Mattioli, E., Gualandi, F., Ferlini, A., Merlini, L., Maraldi, N. M., Bonaldo, P., and Bernardi, P. (2007) *Proc. Natl. Acad. Sci. U. S. A.* **104,** 991–996
23. Pfaff, M., Aumailley, M., Specks, U., Knolle, J., Zerwes, H. G., and Timpl, R. (1993) *Exp. Cell Res.* **206,** 167–176
24. Marinkovich, M. P., Lunstrum, G. P., and Burgeson, R. E. (1992) *J. Biol. Chem.* **267,** 17900–17906
25. Colombatti, A., Ainger, K., and Colizzi, F. (1989) *Matrix* **9,** 177–185
26. Persikov, A. V., Ramshaw, J. A., and Brodsky, B. (2005) *J. Biol. Chem.* **280,** 19343–19349
27. Muzny, D. M., Scherer, S. E., Kaul, R., Wang, J., Yu, J., Sudbrak, R., Buhay, C. J., Chen, R., Cree, A., Ding, Y., Dugan-Rocha, S., Gill, R., Gunaratne, P., Harris, R. A., Hawes, A. C., *et al.* (2006) *Nature* **440,** 1194–1198
28. Soderhall, C., Marenholz, I., Kerscher, T., Ruschendorf, F., Esparza-Gordillo, J., Worm, M., Gruber, C., Mayr, G., Albrecht, M., Rohde, K., Schulz, H., Wahn, U., Hubner, N., and Lee, Y. A. (2007) *PLoS Biol.* **5,** e242
29. Dziadek, M., Kazenwadel, J. S., Hendrey, J. A., Pan, T. C., Zhang, R. Z., and Chu, M. L. (2002) *Matrix Biol.* **21,** 227–241
30. Lamande, S. R., Sigalas, E., Pan, T. C., Chu, M. L., Dziadek, M., Timpl, R., and Bateman, J. F. (1998) *J. Biol. Chem.* **273,** 7423–7430
31. Chu, M. L., Conway, D., Pan, T. C., Baldwin, C., Mann, K., Deutzmann, R., and Timpl, R. (1988) *J. Biol. Chem.* **263,** 18601–18606
32. Knupp, C., and Squire, J. M. (2001) *EMBO J.* **20,** 372–376
33. Lamande, S. R., Morgelin, M., Adams, N. E., Selan, C., and Allen, J. M. (2006) *J. Biol. Chem.* **281,** 16607–16614
34. Perona, J. J., Tsu, C. A., Craik, C. S., and Fletterick, R. J. (1993) *J. Mol. Biol.* **230,** 919–933
35. Bradley, M., Soderhall, C., Luthman, H., Wahlgren, C. F., Kockum, I., and Nordenskjold, M. (2002) *Hum. Mol. Genet.* **11,** 1539–1548
36. Bischof, J. M., Chiang, A. P., Scheetz, T. E., Stone, E. M., Casavant, T. L., Sheffield, V. C., and Braun, T. A. (2006) *Hum. Mutat.* **27,** 545–552
37. Bu, L. M., Bradley, M., Soderhall, C., Wahlgren, C. F., Kockum, I., and Nordenskjold, M. (2006) *Allergy* **61,** 617–621
38. Ober, C., Cox, N. J., Abney, M., Di Rienzo, A., Lander, E. S., Changyaleket, B., Gidley, H., Kurtz, B., Lee, J., Nance, M., Pettersson, A., Prescott, J., Richardson, A., Schlenker, E., Summerhill, E., Willadsen, S., and Parry, R. (1998) *Hum. Mol. Genet.* **7,** 1393–1398
39. Baker, N. L., Morgelin, M., Peat, R., Goemans, N., North, K. N., Bateman, J. F., and Lamande, S. R. (2005) *Hum. Mol. Genet.* **14,** 279–293
40. Lampe, A. K., Dunn, D. M., von Niederhausern, A. C., Hamil, C., Aoyagi, A., Laval, S. H., Marie, S. K., Chu, M. L., Swoboda, K., Muntoni, F., Bonnemann, C. G., Flanigan, K. M., Bushby, K. M. D., and Weiss, R. B. (2005) *J. Med. Genet.* **42,** 108–120
41. Lucioli, S., Giusti, B., Mercuri, E., Vanegas, O. C., Lucarini, L., Pietroni, V., Urtizberea, A., Ben Yaou, R., de Visser, M., van der Kooi, A. J., Bonnemann, C., Iannaccone, S. T., Merlini, L., Bushby, K., Muntoni, F., Bertini, E., Chu, M. L., and Pepe, G. (2005) *Neurology* **64,** 1931–1937
42. Lee, J. O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995) *Cell* **80,** 631–638
43. Perkins, S. J., Smith, K. F., Williams, S. C., Haris, P. I., Chapman, D., and Sim, R. B. (1994) *J. Mol. Biol.* **238,** 104–119
44. Sengle, G., Kobbe, B., Morgelin, M., Paulsson, M., and Wagener, R. (2003) *J. Biol. Chem.* **278,** 50240–50249