ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/12282693

Properties of the collagen type XVII ectodomain - Evidence for N- to C-terminal triple helix folding

Article *in* Journal of Biological Chemistry · February 2001

DOI: 10.1074/jbc.M008709200 · Source: PubMed

CITATIONS
71

READS
147

7 authors, including:

Sayed Kamel Areida
Mansoura University

23 PUBLICATIONS **117** CITATIONS

SEE PROFILE

Dieter Peter Reinhardt
McGill University

177 PUBLICATIONS **8,749** CITATIONS

SEE PROFILE

Holger Notbohm
Universität zu Lübeck

89 PUBLICATIONS **2,854** CITATIONS

SEE PROFILE

All content following this page was uploaded by Sayed Kamel Areida on 02 March 2015.

The user has requested enhancement of the downloaded file.

jbc The Journal of Biological Chemistry
AFFINITY SITES
ASBMB
American Society for Biochemistry and Molecular Biology

PROTEIN STRUCTURE AND FOLDING:

Properties of the Collagen Type XVII Ectodomain: EVIDENCE FOR N- TO C-TERMINAL TRIPLE HELIX FOLDING

Sayed K. Areida, Dieter P. Reinhardt, Peter K. Müller, Peter P. Fietzek, Jutta Köwitz, M. Peter Marinkovich and Holger Notbohm

J. Biol. Chem. 2001, 276:1594-1601.
doi: 10.1074/jbc.M008709200 originally published online October 20, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M008709200

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/276/2/1594.full.html#ref-list-1

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 276, No. 2, Issue of January 12, pp. 1594–1601, 2001
Printed in U.S.A.

Properties of the Collagen Type XVII Ectodomain

EVIDENCE FOR N- TO C-TERMINAL TRIPLE HELIX FOLDING*

Received for publication, September 22, 2000, and in revised form, October 20, 2000
Published, JBC Papers in Press, October 20, 2000, DOI 10.1074/jbc.M008709200

Sayed K. Areida‡§, Dieter P. Reinhardt‡, Peter K. Müller‡¶, Peter P. Fietzek‡, Jutta Köwitz‖, M. Peter Marinkovich, and Holger Notbohm‡ ‡‡**

*From the ‡Medizinische Universität zu Lübeck, Institut für Medizinische Molekularbiologie, Ratzeburger Allee 160, D-23538 Lübeck, Germany, **Dermatology Service, Veterans Affairs Palo Alto Health Care System, Palo Alto, California 94304 and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, California 94305, and ‖Orthopädische Klinik, Universität Rostock, Ulmenstrasse 44, D-18057 Rostock, Germany*

Collagen XVII is a transmembrane component of hemidesmosomal cells with important functions in epithelial-basement membrane interactions. Here we report on properties of the extracellular ectodomain of collagen XVII, which harbors multiple collagenous stretches. We have recombinantly produced subdomains of the collagen XVII ectodomain in a mammalian expression system. rColXVII-A spans the entire ectodomain from δNC16a to NC1, rColXVII-B is similar but lacks the NC1 domain, a small N-terminal polypeptide rColXVII-C encompasses domains δNC16a to C15, and a small C-terminal polypeptide rColXVII-D comprises domains NC6 to NC1. Amino acid analysis of rColXVII-A and -C demonstrated prolyl and lysyl hydroxylation with ratios for hydroxyproline/proline of 0.4 and for hydroxylysine/lysine of 0.5. A small proportion of the hydroxylysyl residues in rColXVII-C (~3.3%) was glycosylated. Limited pepsin and trypsin degradation assays and analyses of circular dichroism spectra clearly demonstrated a triple-helical conformation for rColXVII-A, -B, and -C, whereas the C-terminal rColXVII-D did not adopt a triple-helical fold. These results were further substantiated by electron microscope analyses, which revealed extended molecules for rColXVII-A and -C, whereas rColXVII-D appeared globular. Thermal denaturation experiments revealed melting temperatures of 41 °C (rColXVII-A), 39 °C (rColXVII-B), and 35 °C (rColXVII-C). In summary, our data suggest that triple helix formation in the ectodomain of ColXVII occurs with an N- to C-terminal directionality.

Collagen XVII, also known as the 180-kDa bullous pemphigoid antigen (BP180), is a transmembrane protein that is widely known as a structural component of hemidesmosomes, although structures at cell-tissue interfaces other than hemidesmosomes may also contain collagen XVII (1, 2). Mutations in the collagen XVII gene, *COL17A1*, lead to junctional epidermolysis bullosa, a hereditary blistering skin disease with epidermal detachment from the basement membrane (3).

The cDNA sequence of collagen XVII encodes a type II integral transmembrane protein of 1497 amino acid residues (4). It consists of an intracellular domain of 466 residues, a transmembrane domain of 23 residues, and an extracellular collagenous domain of 1008 amino acids with multiple non-collagenous interruptions. The length of the individual collagenous regions varies from 14 to 242 amino acid residues (4). Collagen XVII exists in two molecular forms, *i.e.* as a full-length transmembrane homotrimer of three 180-kDa α1(XVII) chains and as a 120-kDa soluble form. The latter corresponds to the extracellular domain and is presumably released from the cell surface by furin-mediated proteolytic processing (5). In some instances, an even shorter fragment with ~90–100 kDa has been observed (6).

Some information about the molecular shape of collagen XVII under physiological conditions can be deduced from rotary shadowing electron microscopy of collagen XVII from bovine cell lines or from recombinant fragments. These studies revealed asymmetric molecules with an elongated shape and a globular, ball-like structure at one end (6, 7). A 90-kDa pepsin/trypsin fragment of collagen XVII in detergent extracts of keratinocytes was resistant to further trypsin digestion at physiological temperatures, therefore suggesting that it was of triple-helical structure (5).

The entire collagen XVII as well as the C15 collagenous domain could be recombinantly produced in protease-resistant conformation, apparently not requiring propeptides (7). Triple helix formation of fiber-forming collagens was defined as a two-step process initiated by annealing and aggregation of the C-propeptides followed by C- to N-terminal folding of the three α-chains (8). A higher content of hydroxyproline residues in the C-terminal region of the amino acid sequence of fibrillar collagens has been considered as evidence that the C terminus is the locus of initiation of triple helix formation (8).

Opposite to this, there is also experimental evidence that C-propeptides can be substituted for by other peptide sequences or can be deleted without apparent harm for helix formation (9–11). In fibril-associated collagens with interrupted triple helices, C-propeptides are completely missing, and even if nonhelical ends at the C terminus are deleted in recombinantly produced collagenous molecules, as shown for collagen XII, correct triple helix formation takes place (12).

The present study was designed to address the role of subdomains of collagen XVII ectodomain in triple helix formation and stability, in particular to delineate information on the directionality of the assembly process. With a variety of recombinant polypeptides expressed in a mammalian system, we

* This work was supported by Deutsche Forschungsgemeinschaft Grants DFG-No-147-6 (to H. N.), and Re 1021/4–1 (to D. P. R.), and National Institutes of Health Grant P01-AR 44012 (to M. P. M.). 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.

§ Supported by a scholarship from the Zoological Department, Faculty of Science, University of Mansura, Mansura, Egypt.

¶ Recipient of a Volkswagen Foundation stipend for his stay on sabbatical leave in Stanford.

‡‡ To whom correspondence should be addressed. Tel.: 49-451-500-4083; Fax: 49-451-500-3637; E-mail: notbohm@molbio.mu-luebeck.de.

This paper is available on line at http://www.jbc.org

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

obtained structural information by circular dichroism studies, protease sensitivity assays, and electron microscope studies that strongly suggest triple helix formation from the N to the C terminus of collagen XVII.

EXPERIMENTAL PROCEDURES

Construction of Expression Plasmids for Subdomains of the Collagen XVII Ectodomain—Human cDNA coding for collagen XVII (4) was used as the template for amplification by the polymerase chain reaction with appropriate primers. The numbering of nucleotides in cDNAs are according to Giudice *et al.* (Ref. 4, GenBank™ accession number M91669), and the deduced protein sequence was counted from the methionine start codon (position 36 in M91669). All fragments were digested by appropriate restriction enzymes and were cloned into an episomal expression vector pCEP-Pu containing the signal peptide sequence of BM-40 and a puromycin resistance gene (13). Cloning of inserts into pCEP-Pu via a *Nhe*I restriction site resulted in secreted polypeptides with four additional N-terminal amino acid residues (APLA) preceding the expressed sequence (13). Sequences and correct in-frame insertions of all constructs were verified by DNA sequencing (Medigenomix).

An expression plasmid for rColXVII-A corresponding to the extracellular domain of collagen XVII (amino acid residues 527–1497) was constructed by amplification of the cDNA with the sense primer 5′-TTCGCTAGCTATGGCACCCGCGGCGGGAGCAGAC-3′ (nucleotides 1684 to 1707) and the antisense primer 5′-ACGCGTCGACTCACGGCT-TGACAGCAATACTTCTTC-3′ (nucleotides 4574 to 4596). The *Nhe*I-*Sal*I fragment of the product was subcloned into pCEP-Pu, resulting in plasmid pCEP-rColXVII-A. This expression plasmid generates a 971-amino acid residue protein with the N terminus within the NC16a domain and the C terminus at the cognate end of the ectodomain.

The expression plasmid for rColXVII-B (amino acid residues 527–1482) was constructed by amplification of the cDNA with sense primer 5′-TTCGCTAGCTATGGCACCCGCGGCGGGAGCAGAC-3′ (nucleotides 1684 to 1707) and antisense primer 5′-ACGCGTCGACTCATTG-GTCACCTTTGTCTCCTTTTTCTC-3′ (nucleotides 4526 to 4551). The *Nhe*I-*Sal*I-restricted fragment from the amplification product was ligated into pCEP-Pu, resulting in plasmid pCEP-rColXVII-B. This plasmid encodes 956 amino acid residues and is identical to rColXVII-A except for a 15-amino acid deletion (NC1 domain) at the C-terminal end.

To prepare the expression plasmid for rColXVII-C (amino acid residues 527–808), the cDNA was amplified with sense primer 5′-TTCGCTAGCTGAGGAGGTGAGGAAGCTG-3′ (nucleotides 1573 to 1590) and the antisense primer 5′-ACGCGTCGACTCAGATCTTGCCT-GGAG-3′ (nucleotides 2516 to 2529). The *Nhe*I-*Sal*I fragment from this product was ligated into pCEP-Pu, resulting in plasmid pCEP-rColXVII-C. This plasmid encodes a 282-amino acid residue polypeptide comprising 40 residues of the NC16a domain (δNC16a) and the C15 domain.

The expression plasmid for rColXVII-D (amino acid residues 1188–1497) was generated by cDNA amplification using the sense primer 5′-TTCGCTAGCTCCAGGCAATGTGTGGTCCAGCATC-3′ (nucleotides 3670 to 3693) and the antisense primer 5′-ACGCGTCGACTCACGGCT-TGACAGCAATACTTCTTC-3′ (nucleotides 4574 to 4596). The *Nhe*I-*Sal*I fragment was ligated into pCEP-Pu, resulting in plasmid pCEP-rColXVII-D. This plasmid encodes a 310-amino acid polypeptide spanning from the NC1 to the NC6 domain.

Generation of Recombinant Cell Clones and Production of Conditioned Medium—Human embryonic kidney 293-EBNA cells (Invitrogen) were cultivated in Dulbecco's modified Eagle's medium (Life Technologies) containing 10% fetal calf serum, 0.25 mg/ml G418 (Calbiochem), 0.1 mg/ml penicillin/streptomycin (Biochrom KG), and 2 mM L-glutamine (Biochrom KG). The pCEP-rColXVII plasmids (25 μg) were separately transfected into 293-EBNA cells (one million cells/10 cm^2 culture dish) using a calcium phosphate precipitation method (14). After a selection with 0.5 μg/ml puromycin (Calbiochem), the transfected cells were grown to confluence, washed twice with phosphate-buffered saline, and switched to serum-free medium containing 50 μg/ml ascorbic acid (Sigma), freshly prepared. The media were collected every 24 h, cooled, centrifuged to remove cellular debris, supplemented with 1 mM protease inhibitor Pefablock (Roth), and frozen at -80 °C until further use.

Purification of Recombinant Collagen Type XVII Subdomains—The serum-free media (2–3 liters) containing recombinant fragments of rColXVII-A and rColXVII-B were concentrated to ~80 ml by ultrafiltration and dialyzed against 20 mM Tris-HCl, pH 8.8, at 4 °C. The material was passed over an anion exchange column (HiTrapQ, 5 ml; Amersham Pharmacia Biotech) equilibrated in the same buffer and subsequently eluted with a linear 0–0.4 M NaCl gradient. Fractions containing the recombinant fragments were concentrated by ultrafiltration to ~0.5 ml and passed over a Superose 6 column (30 ml; Amersham Pharmacia Biotech) equilibrated in 50 mM Tris-HCl, pH 8.6, 150 mM NaCl at a flow rate of 0.3 ml/min. Fractions containing recombinant fragments were detected by SDS gel electrophoresis and Western blotting.

The serum-free cell culture media (2–3 liters) containing recombinant fragments of rColXVII-C and rColXVII-D were concentrated to ~70 ml, dialyzed against 50 mM sodium acetate, pH 4.8, at 4 °C, passed over a cation exchange column (HiTrapS; 5 ml; Amersham Pharmacia Biotech) equilibrated in the same buffer, and eluted with a linear 0–0.4 M NaCl gradient. Fractions containing the recombinant fragments were concentrated to ~0.5 ml and passed over a Superose 12 column (30 ml; Amersham Pharmacia Biotech) equilibrated in 50 mM sodium acetate, pH 4.8, 200 mM NaCl. Fractions were collected, and recombinant fragments were identified as described above.

Electrophoresis and Western Blot Analysis—SDS gel electrophoresis (15) and Western blot analysis (16) were performed according to standard procedures. 6% polyacrylamide gels were used for rColXVII-A and -B, and 10% polyacrylamide gels were used for rColXVII-C and -D. Polyclonal rabbit antiserum NC16a was used 1:1000 diluted for rColXVII-A, -B, and -C (17), and polyclonal chicken antiserum Col17ecto-1 (5) was used 1:20-diluted for rColXVII-D. The antisera were a generous gift of Dr. L. Bruckner-Tuderman. Goat-anti rabbit (1:1000; Dako) or goat anti-chicken (1:100; Dako) antibodies conjugated to horseradish peroxidase were used as secondary antibodies.

Circular Dichroism Measurements—Purified recombinant rColXVII polypeptides were adjusted to a concentration of 20 μg/ml and dialyzed against 0.05% acetic acid. Far UV circular dichroism spectra (190–260 nm) were recorded in a 1-cm quartz cuvette on a Jasco J-715 A spectropolarimeter equipped with a temperature controller. The molar ellipticities were calculated on the basis of a mean residue molar mass of 96 g/mol. Thermal transition curves were recorded at a fixed wavelength (221 nm) by raising the temperature linearly at a rate of 30 °C/h using a Gilford temperature programmer. No deviation from the melting profiles were observed with extended temperature gradients. The degree of triple helicity at various temperatures was calculated by setting the 221-nm signal at 20 °C to 1 (maximal amount of triple helicity) and at 45 °C to zero (completely denatured collagen).

Low Angle Rotary Shadowing Electron Microscopy—rColXVII polypeptides were dialyzed against a solution of 50% glycerol in 0.05% acetic acid for 16 h at 4 °C. Samples were sprayed onto freshly cleaved mica using an air brush. The droplets on the mica were dried at room temperature at 10^{-6} mm Hg for 12 h in a vacuum coater (Edwards 306). The dried specimens were rotary-shadowed with platinum using an electron gun positioned at 6° to the mica surface and then coated with a film of carbon generated by an electron gun positioned at 90° to the mica surface. The replica were floated on distilled water and collected on formvar-coated grids. The replicas were examined on a Zeiss 109 transmission electron microscope. Length and diameter of molecules and aggregates were determined with the Scion Image program (Scion Image Corporation).

Analysis of Hydroxylation and Glycosylation—For amino acid analyses, purified recombinant rColXVII polypeptides were hydrolyzed with 6 N HCl for 24 h at 110 °C. Samples were analyzed on an amino acid analyzer (Biochrom 20, Amersham Pharmacia Biotech) using ninhydrin for post-column color development.

For determination of glucosylgalactosylhydroxylysine and galactosylhydroxylysine samples, up to 1 mg of purified rColXVII-C was hydrolyzed in 1 ml of 2 N KOH for 24 h at 110 °C. After hydrolysis, 100 μl of glacial acetic acid and 150 μl of 70% perchloric acid were added, mixed, and centrifuged for 10 min at 14,000 rpm. The supernatant was decanted and lyophilized. Lyophilized samples were re-dissolved in 1 ml of water and passed over a CF1 (Whatman) column to remove the majority of amino acids. Eluates were lyophilized and analyzed on an amino acid analyzer (Biochrom 20, Amersham Pharmacia Biotech).

N-terminal Sequence Analysis—Purified native or pepsin-digested rColXVII-C and -D was subjected to SDS gel electrophoresis and transferred onto a Mini Pro Blott membrane (Applied Biosystems). Protein bands were visualized by Coomassie Blue staining and identified by comparison with immunoblotted material. Relevant protein bands were excised and loaded directly onto a Procise 494 protein sequencer (Applied Biosystems) for N-terminal sequencing.

Enzyme Digests—For assessment of the domain structure and stability of collagen XVII, purified recombinant rColXVII polypeptides

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

were subjected to treatment by various enzymes. Collagenase treatment was used to release collagenous peptides, and pepsin treatment was used to remove protease-sensitive non-collagenous regions. Trypsin digestion was used to determine the melting temperature of the rColXVII fragments A and C.

Digestion of purified polypeptides with highly purified bacterial collagenase (Sigma) was performed with 40 units/ml enzyme in 0.2 M NH_4HCO_3, freshly prepared, at 37 °C for 2 h (18). For pepsin digestion, 100 μl of rColXVII polypeptides (~25 μg) were acidified by dialysis against 0.05% acetic acid and incubated with 1 μg/ml pepsin (Roche Diagnostics) at 4 °C for 24 h (19). After neutralization with saturated Tris solution, samples were separated by SDS gel electrophoresis followed by Western blotting using anti-NC16a and anti Col17ecto-1 antisera.

For testing the triple-helical conformation with trypsin (20), the purified samples (~20 μg/ml) were dialyzed against 100 mM Tris-HCl, pH 7.4, 0.4 M NaCl and preheated for 5 min at each desired temperature between 20 and 46 °C (2 °C steps). An aliquot of 10 μl was removed, cooled quickly to 20 °C, and treated with 10 μl of trypsin (1 mg/ml; Sigma) for 2 min. Reactions were stopped by adding 10 μl of soybean trypsin inhibitor (5 mg/ml; Sigma) to each individual sample. The incubated samples were analyzed by SDS gel electrophoresis followed by Western blotting with specific antibodies. Intensities of the immunoblotted bands were quantified with Gel-Pro-Analyzer (Media Cybernetics).

RESULTS

Production and Characterization of Recombinant Col XVII Subdomains—To investigate triple-helix formation and stability of the collagen type XVII ectodomain, we have produced several recombinant subfragments of the ectodomain in mammalian 293 cells. These subdomains included the full-length ectodomain rColXVII-A (positions 527–1497), a C-terminal truncated form thereof, rColXVII-B (positions 527–1482), the largest collagenous stretch, C15 (positions 527–808), and a short C-terminal subdomain rColXVII-D (positions 1188–1497) spanning the NC6 to the NC1 domain. These recombinant fragments are schematically summarized in Fig. 1. Each of the constructs were designed with a heterologous signal peptide to enable secretion into the cell culture medium and proper post-translational modifications. The recombinant cell clones produced each subdomain in quantities of 1–2 μg of protein/ml medium/day. Ion exchange chromatography followed by gel filtration chromatography resulted in highly purified recombinant subdomains with apparent molecular masses of 101 kDa (rColXVII-A), 96 kDa (rColXVII-B), 30 kDa (rColXVII-C), and 33 kDa (rColXVII-D), which corresponded well with the calculated values (Fig. 2). Interestingly, the collagenous domain fragment rColXVII-C migrates as a broad band, similar to what has been published recently for a corresponding construct (21). No multimers have been observed after electrophoresis by SDS gel electrophoresis (Fig. 2). N-terminal sequencing of rColXVII-C, which N-terminally starts at the same position as rColXVII-A and -B, demonstrated the expected N-terminal sequence (APLAMA) but also some minor N-terminal truncations of 2–5 amino acid residues (Table I). Sequencing of the C-terminal construct rColXVII-D resulted in a sequence (APGNVWSSIS) that was three amino acid residues shorter than expected (Table I). These results indicate that the sequence APLA, which represents an artificial sequence due to the expression strategy (see "Experimental Procedures"), is not stable in the recombinant constructs. The authentic collagen type XVII sequences, however, were stable.

The rColXVII-A and rColXVII-C polypeptide were further analyzed by amino acid analysis. Within the limits of error, the amino acid compositions determined correlated well with the expected compositions calculated from the cDNA (Table II). For rColXVII-A, about 30% of the total prolyl and 34% of the total lysyl residues were hydroxylated. For rColXVII-C, about 29% of

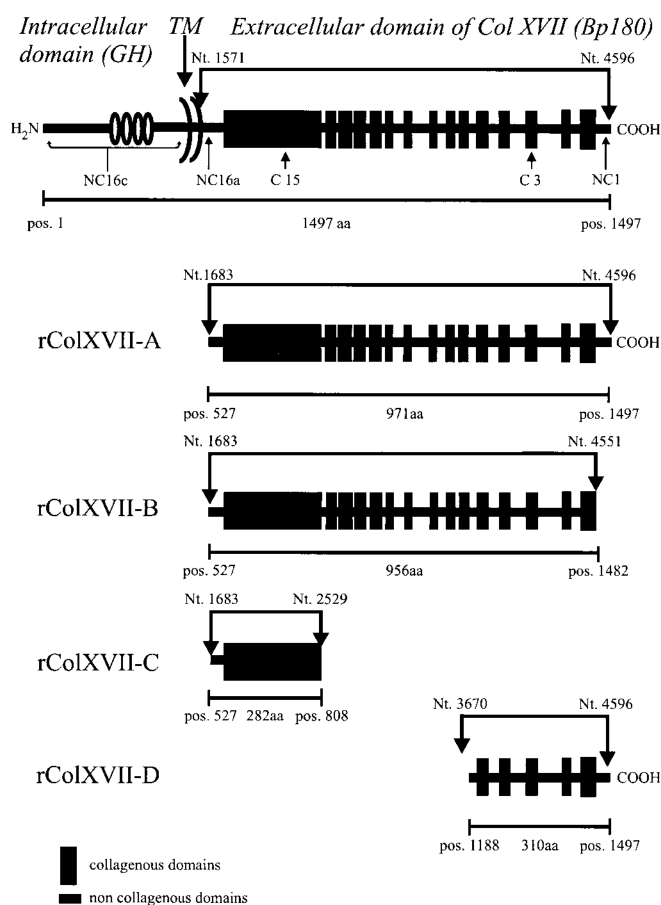

FIG. 1. **Schematic representation of recombinant subdomain constructs of human collagen XVII (BP 180).** The extracellular domain consists of a series of 15 collagenous domains (C 1 through C 15, *solid vertical boxes*) and stretches of non-collagenous domains (NC 1 through NC16a, *horizontal bars*). The rColXVII-A construct extends from nucleotide 1683 to 4596 (amino acid positions 527–1497). rColXVII-B construct extends from nucleotide 1683 to 4551 (amino acid positions 527–1482). rColXVII-C construct extends from nucleotide 1683 to 2529 (amino acid positions 527–808). rColXVII-D construct extends from nucleotide 3670 to 4596 (amino acid positions 1188–1497). The total numbers of amino acid residues for each of the constructs are indicated. *GH*, globular head. *TM*, transmembrane region; *Nt*, nucleotide; *aa*, amino acids.

the total prolyl and 36% of the total lysyl residues were hydroxylated (Table II). Some of these residues in rColXVII-C were also further modified by the attachment of a small but significant number of mono- and disaccharides (0.01 monosaccharide and 0.17 disaccharide per molecule; Table II).

Pepsin and Collagenase Degradation of Recombinant Col XVII Subdomains—Limited enzymatic digestion with pepsin converted rColXVII-A into a ~30-kDa fragment via an intermediate protein of the apparent molecular mass of 58 kDa (Fig. 3, *panel A*, *lane 2*), whereas digestion with bacterial collagenase completely degraded this recombinant ectodomain under the experimental conditions. Similar data were obtained for the C-terminally truncated rColXVII-B (Fig. 3, *panel B*). The 30-kDa pepsin fragments corresponded to the size and the immunoreactivity of rColXVII-C (collagen XVII domain C15), indicating that this domain adopts a stable triple-helical conformation in rColXVII-A and -B and, thus, resists degradation by pepsin. Recombinant rColXVII-C itself was also resistant to limited degradation with pepsin, whereby a slightly (~ 3 kDa) faster electrophoretic mobility of the pepsin-treated material was observed (Fig. 3, *panel C*). N-terminal sequencing of the pepsin-treated rColXVII-C revealed a N-terminally trunca-

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

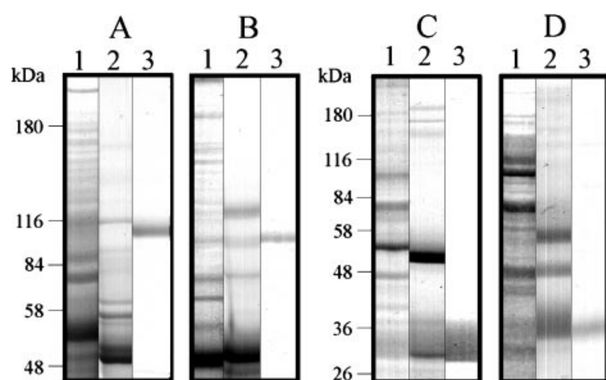

FIG. 2. **Purification of recombinant extracellular subdomains of human collagen type XVII.** rColXVII polypeptides were purified, separated by SDS-gel electrophoresis, and stained with Coomassie Blue. *A* and *B*, rColXVII-A and rColXVII-B, respectively; control medium from nontransfected 293-EBNA cells (*lane 1*), protein pattern after HiTrapQ anion exchange chromatography (*lane 2*), and Superose 6 gel filtration chromatography (*lane 3*). *C* and *D*, rColXVII-C and rColXVII-D, respectively; control medium from nontransfected 293-EBNA cells (*lane 1*), protein pattern following HiTrapS cation exchange chromatography (*lane 2*), and Superose 12 gel filtration chromatography (*lane 3*). The relatively broad protein band in *panel C* (*lane 3*) is probably due to heterogeneity in glycosylation or in the amino acid sequence at the N-terminal end. Positions of globular marker proteins are indicated in kDa.

tion of 32 amino acid residues, corresponding to parts of the non-collagenous δNC16a domain (Table I). Accordingly, the Gly-X-Y repeats of rColXVII-C were not degraded by pepsin. These results suggest a triple-helical conformation for rColXVII-C and indicate that more C-terminally located domains are not required for triple helix formation of domain C15. Treatment of the C-terminal rColXVII-D with pepsin resulted in a complete degradation, indicating the absence of a stabilizing triple-helical conformation (Fig. 3, *panel D*). Controls with bacterial collagenase demonstrated complete degradation of all recombinant polypeptides (Fig. 3, *lanes 3*).

Ultrastructural Analysis of Recombinant rColXVII Subdomains—To obtain information of the molecular shape, the recombinant rColXVII subdomains were analyzed by electron microscopy after rotary shadowing (Fig. 4). The full-length ectodomain rColXVII-A displayed a long extended shape with a globular portion on one end (Fig. 4*A*). Measurement of the lengths of individual particles revealed 3 groups of about 60–70 nm, 130–140 nm, and 240–250 nm. A length of 130–140 nm corresponds well with expected values for the ectodomain (22). A length of 240–250 nm may represent dimers that occur through lateral alignment, whereas shorter fragments may derive from proteolytic cleavage of the polypeptide. The globular domain at one end is about 11–12 nm in diameter. The widths of the rotary-shadowed molecules (~ 4 nm) are consistent with the width of triple-helical collagen molecules, accounting for about 2 nm for the platinum coat. The molecules often adopt relatively rigid conformations including some wave-like regions but without the presence of sharp kinks or bends.

Rotary-shadowed molecules of rColXVII-C display an extended thread-like shape with a uniform length distribution of 63 ± 5 nm (Fig. 4*B*). The extended shape and the width of the molecules clearly indicate a triple-helical conformation of rColXVII-C. No globular regions nor obvious kinks were observed. Since this construct starts at the same amino position as rColXVII-A, we conclude that the globular domain in rColXVII-A represents the C-terminal end of the molecule.

The C-terminal rColXVII-D always resulted in globular particles of about 10–12 nm in diameter, with some resemblance to the globular domain at the C-terminal end of rColXVII-A (Fig. 4*C*). No extended regions, suggestive for triple-helical collagenous domains, were observed. Since rColXVII-D contains five collagenous regions (C1-C5), we concluded that this recombinant construct was not able to adopt triple-helical conformations and instead assembled into a globular pepsin-sensitive aggregate (see also Fig. 3*D*, *lane 2*).

Conformation and Stability of the Recombinant Collagen XVII Fragments—To further investigate structural aspects, the recombinant subdomains rColXII-A, -B, and -C were analyzed by spectral and thermal circular dichroism analyses. For rColXVII-C,a spectrum typical for a collagen triple helix with a maximum at 221 nm was observed (Fig. 5*A*). Purified molecules that include non-collagenous domains (rColXII-A and -B) demonstrated shoulders at 221 nm (Fig. 5*A*). These shoulders did not reach values of positive ellipticities typical for triple helices. However, when rColXVII-A (Fig. 5*A*) or -B (not shown) were treated with pepsin to remove the non-collagenous regions before analysis, typical maxima at 221 nm were observed. These data demonstrate that pepsin-resistant portions of rColXVII-A or -B adopt a triple-helical conformation. We concluded that in intact rColXVII-A or -B the ellipticities originating from non-collagenous regions interfered with those originating from collagenous triple-helical regions to produce shoulders instead of maxima at 221 nm.

Thermal denaturation profiles of the recombinant fragments were recorded as a decrease of ellipticity at 221 nm (Fig. 5*B*). Polypeptide rColXVII-C demonstrated a melting temperature of about 35 °C and a melting range from 30 to 40 °C. Full-length ectodomain rColXVII-A melted at about 41 °C, with a melting range of 35 °C to 46 °C. After treatment of rColXVII-A with pepsin, a biphasic melting curve was observed. The first phase of this profile resembles the denaturation profile of rColXVII-C and represents likely the intact C15 triple-helical domain in the proteolytic degradation mixture. The second phase demonstrated a somewhat higher melting temperature of about 38 °C and probably represents domain C15 plus additional C-terminal regions corresponding to the 58-kDa fragment obtained after limited pepsin degradation of rColXVII-A (Fig. 3*A*, *lane 2*). The denaturation profile of rColXVII-B was, as expected, very similar to the results obtained with rColXVII-A, with a range of melting temperature between 34 and 46 °C. After pepsin degradation of rColXVII-B, a biphasic melting curve was observed, and the melting temperature decreased from 39 to 34 °C (data not shown).

The results obtained by circular dichroism analysis were further confirmed by using a temperature-dependent trypsin digestion assay (20). In this test, denaturation of the triple-helical regions of rColXVII-A and -C was measured at various temperatures by monitoring the degradation of the protein by trypsin. The results of this assay are shown in Fig. 6. Melting temperatures of about 43 °C for rColXVII-A (Fig. 6*A*) and of about 33 °C for rColXVII-C (Fig. 6*B*) are in good agreement with data obtained by circular dichroism (Fig. 5*B*). The rColXVII-A is not degraded by trypsin at temperatures below ~35 °C, although many non-collagenous domains containing arginine and lysine residues are present in this polypeptide. This result indicates that the non-collagenous domains are protected against proteolytic attack by trypsin under the conditions applied.

DISCUSSION

Collagen XVII is a transmembrane component of hemidesmosomes and other cell-matrix interfaces. Although the function of the collagen XVII intracellular region has been the focus of intensive analyses (23–25), little is known about structure and function of the ectodomain. Here we report on the struc-

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

TABLE I
N-terminal sequence analysis of rColXVII-C and -D

The sequences derived from the cDNA are compared with results from N-terminal sequencing of constructs rCoIXVII-C and -D or with pepsin-degraded rColXVII-C. ↓ indicates the cleavage site between the signal peptide and the mature polypeptide. The amino acid residues APLA at the N terminus of each construct are a result of the cloning strategy. The first authentic amino acid residues from the α1(XVII) polypeptide and their corresponding positions are indicated in bold face. Gly-X-Y repeats are underlined.

Derived from:	Amino acid sequence
cDNA of rColXVII-C	A ↓ A P L A **M** A P A A G A D L D K I G L H S D S Q E E L W M F V R K K L M M E Q E N G N L R <u>G S P G P K...</u>
	527
rColXVII-C	(1) A P A A G A D L D K........ Major sequences (1–2)
	(2) A P L A M A.........................
	(3) L A M A P A..................... Minor sequences (3–5)
	(4) A M A P A A....................
	(5) M A P A......................
rColXVII-C, pepsin-digested	....E Q E N G N L R <u>G S.......</u>
cDNA of rColXVII-D	A ↓ A P L A **P** G N V W S S I S V E D L S
	1188
rColXVII-D	A P G N V W S S I S..........

TABLE II
Amino acid and glycosylation analysis of rColXVII-A and -C

Residue	rColXVII-A residues/molecule		rColXVII-C residues/molecule	
	Determined	Calculated (cDNA)	Determined	Calculated (cDNA)
HypHyp	36.7	Not possible	14.3	Not possible
Asx	54.4	35 (Asp) 11 (Asn)	13.3	10 (Asp) 2 (Asn)
Thr	31.3	26	5.6	5
Ser	99.5	99	11.0	11
Glx	77.8	40 (Glu) 38 (Gln)	30.2	16 (Glu) 12 (Gln)
Pro	86.5	179	35.1	55
Gly	257.8	229	84.2	84
Ala	56.2	48	12.9	14
Cys	0.0	0	0.0	0
Val	36.3	35	9.1	9
Met	15.2	23	7.7	11
Ile	27.8	27	5.3	5
Leu	63.3	61	16.1	15
Tyr	16.0	24	1.6	0
Phe	21.0	18	3.4	2
His	19.5	12	4.5	4
Hyl	11.7	Not possible	5.5	Not possible
Lys	22.7	27	9.7	17
Arg	37.1	42	13.3	13
Hyp/Pro	0.4	Not possible	0.4	Not possible
Hyl/Lys	0.5	Not possible	0.5	Not possible
Trp	Not determined	2	Not determined	1
Monosaccharide	Not determined	Not possible	0.01	Not possible
Disaccharide	Not determined	Not possible	0.17	Not possible

tural analysis of recombinantly expressed fragments of the collagen type XVII ectodomain.

We have recombinantly expressed four constructs composing various regions of the collagen XVII ectodomain: the full-length ectodomain (rColXVII-A), the ectodomain lacking the short C-terminal NC1 domain (rColXVII-B), the longest collagenous region C15 plus 40 N-terminal amino acid residues from the NC16a domain (rColXVII-C), and a C-terminal fragment spanning NC1 to NC6 (rColXVII-D). These recombinant polypeptides were expressed in mammalian 293 cells, a system that previously has been reported to produce correctly folded recombinant collagen XVII constructs (21). All recombinant constructs were designed with a heterologous sequence for a signal peptide to ensure that the recombinant polypeptides pass through the endoplasmic reticulum for post-translational modification. The secreted polypeptides were purified to homogeneity from the cell culture medium by conventional chromatography procedures.

The structure of the recombinant collagen XVII fragments were extensively studied. Analysis of circular dichroism spectra and limited enzymatic digestion with pepsin and trypsin clearly demonstrate that the large fragments rColXVII-A and -B and the much shorter fragment rColXVII-C adopt triple-helical conformations. Based on hydrodynamic properties of the ectodomain, recombinantly expressed in a transient system with COS cells, the ectodomain of collagen XVII has been described as elongated, and a trimeric assembly of α-chains was expected from sedimentation analysis (7). Furthermore, limited trypsin digestion of authentic collagen XVII suggested a triple-helical conformation of the ectodomain (5). In our studies, we analyzed for the first time the recombinant ectodomain of collagen XVII, represented by rColXVII-A and -B, by circular dichroism. The typical maxima for triple helices at 221 nm were less prominent as compared with "pure" collagenous proteins. However, when rColXVII-A and -B were digested with pepsin, the typical maxima at 221 nm could be observed. These results indicate that the signals from triple-helical regions and non-triple-helical regions interfere, providing an explanation for the reduced maxima at 221 nm. Thermal transition curves recorded at 221 nm showed a melting temperature of 41 °C for rColXVII-A, demonstrating a stable triple-helical conformation at physiological temperatures. These data are in good agree-

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

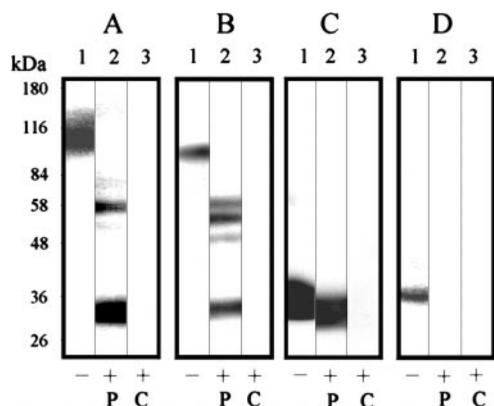

FIG. 3. **Proteolytic fragmentation of the recombinant extracellular subdomains of collagen XVII.** Purified rColXVII polypeptides (*lane 1*) were digested with either 1 μg/ml pepsin (*P*; *lane 2*) at 4 °C for 24 h or with bacterial collagenase (*C*; *lane 3*) at 37 °C for 2 h. Detection of the proteolytic fragments was performed with specific antibodies after Western blotting (see "Experimental Procedures"). The *panel labels* correspond to the individual recombinant subdomains rColXVII-A to -D. The position of marker proteins are indicated in kDa.

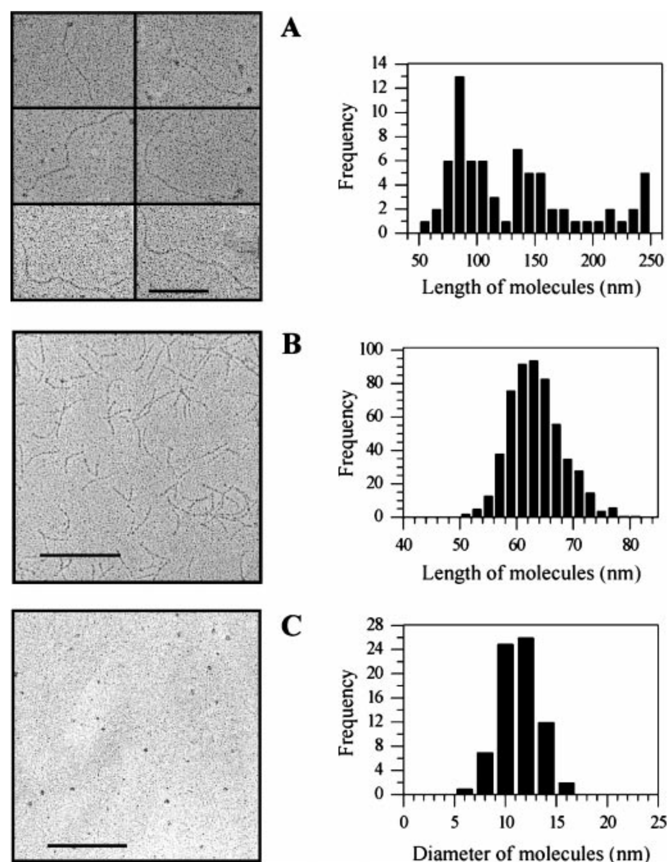

FIG. 4. **Shape and length of recombinant polypeptides.** Electron microscope images of rotary shadowed molecules of rColXVII-A (*A*), rColXVII-C (*B*), and rColXVII-D (*C*). Histograms of measured lengths (rColXVII-A and -C) or diameter (rColXVII-D) of the recombinant polypeptides are shown to the right of each micrograph. Measurements are plotted as numbers of measurements in 10-nm windows (rColXVII-A) or 2-nm windows (rColXVII-C and -D). *Bars* represent 150 nm.

ment with data obtained by a trypsin protection assay of the collagen XVII ectodomain from *ex vivo* preparations (5). Since the C-terminally truncated version rColXVII-B of the recombinant ectodomain displayed very similar circular dichroism spectra and transition profiles, we conclude that the NC1 domain (position 1483–1497) at the C-terminal end does not have

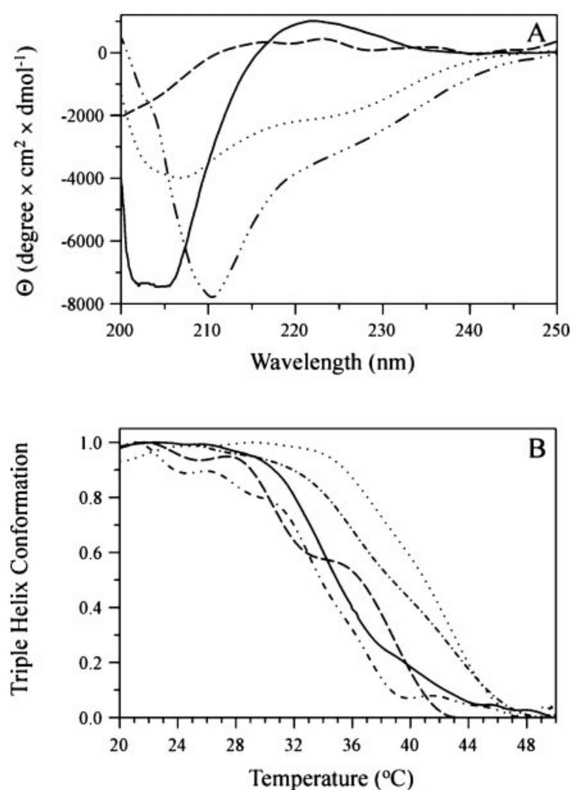

FIG. 5. **Far UV circular dichroism spectra of recombinant collagen XVII subdomains.** *A*, circular dichroism spectra of rColXVII-A (*dotted line*), rColXVII-A treated with pepsin (*dashed line*), rColXVII-C (*solid line*), and rColXVII-D (*dots* and *dashes*). Ellipticity θ is plotted as a function of wavelength. *B*, thermal denaturation profiles of rColXVII polypeptides determined at 221 nm as a function of temperature. The signal at 20 °C was set to 1 and represents the maximal amount of triple helicity in the polypeptides, and the signal at 45 °C was set to 0 and represents completely denatured triple-helical regions. rColXVII-A (*dotted line*), rColXVII-A treated with pepsin (*dashed line*), rColXVII-B (*dash-dot-dash line*), rColXVII-C (*solid line*), rColXVII-C treated with pepsin (*dash-dot-dot-dash line*).

a major impact on the stability of the ectodomain. Furthermore, these results indicate that the NC1 domain is apparently not necessary for the formation of the triple-helical conformation. Comparison of thermal denaturation profiles of rColXVII-A with -C revealed a higher (6 °C) melting temperature for rColXVII-A. These results suggest that at least some of the collagenous domains C1-C14 adopt triple-helical conformations (see also below). Even the non-collagenous domains interspersed between C1-C14 may form relatively tight structures or are stabilized by the contiguous collagenous domains, since rColXVII-A was completely resistant to trypsin up to 35 °C, although a fair number of basic residues resides in this region.

Rotary-shadowed images of rColXVII-A demonstrated extended molecules with a globular domain at one end. The molecules demonstrated several regions of flexibility but no kinks. These results are in agreement with data from sedimentation studies suggesting more flexibility than a rigid rod (7). Electron micrographs from rColXVII-C, the long collagenous stretch, also revealed a long extended shape of 63 nm. Although the molecules appeared relatively rigid, some curved molecules were observed. Based on the absence of a globular domain in rColXVII-C, we can now unambiguously assign the globular domain in rColXVII-A to the C-terminal end, since both fragments start at the same N-terminal position. This C-terminal globular domain was either not present or only occasionally present in other studies (6). In this regard it is interesting that rColXVII-D, corresponding to the C-terminal end, ultrastruc-

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

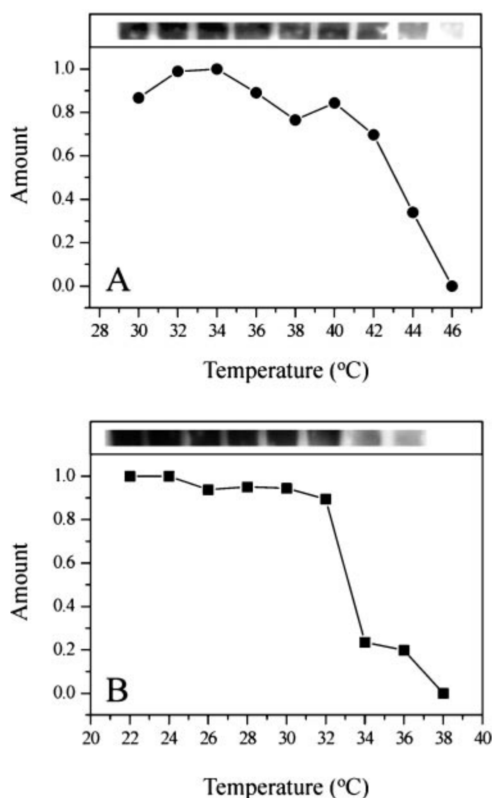

FIG. 6. **Temperature-dependent trypsin digestion assay of rCol XVII-A (*A*) and rCol XVII-C (*B*).** Samples were pre-heated to discrete temperatures in 2 °C steps from 20 to 46 °C and then incubated with trypsin. Degradation was monitored by Western blotting after gel electrophoresis (*upper panels*). Note that rColXII-A and -C were only degraded at elevated temperatures. The intensity of the bands were quantified by densitometry. The protein amounts in arbitrary units are plotted as a function of temperature (*lower panels*).

turally appeared as a globular module of about 10–12 nm in diameter that resembled the globular domain of rColXVII-A. These results are suggestive that this globular domain represents the fold of the five C-terminal collagenous domains C1-C5 interrupted by six non-collagenous regions. However, no evidence for a triple-helical conformation of rColXVII-D was obtained by circular dichroism analysis or by enzyme resistance experiments. These data may indicate that fragment rColXVII-D is not able to adopt its native conformation and rather aggregates to a globular structure.

For rColXVII-C, comprising 81 Gly-X-Y (including one Gly-Ser-Gly) repeats plus 40 additional N-terminal amino acid residues from the NC16a domain, clear evidence for triple-helical conformation was obtained by circular dichroism analysis and limited proteolytic degradation. Thermal denaturation profiles demonstrated a melting temperature of 35 °C, which did not change upon incubation with pepsin. Recently, Tasanen and co-workers (21) recombinantly expressed a similar fragment composing the C15 domain (residues 567–808) in 293-EBNA cells. This fragment lacks the 40 residues originating from the NC16a domain but is otherwise identical to rColXVII-C. Although this recombinant fragment contained a higher amount of hydroxylated proline compared with rColXVII-C (74% of the eligible prolyl residues *versus* 53% in rColXVII-C), the transition curve revealed a significant lower melting temperature (26.5 °C; Ref. 21). This comparison suggest that the 40 additional residues (positions 527–566) N-terminal to the C15 domain contribute somehow to the stability of the C15 domain, even after these residues are removed after pepsin treatment. We suggest that the three α-chains in the NC16a domain interact with each other to allow the formation of a proper stagger within the C15 domain that otherwise would not form. This would provide an explanation of why pepsin-trimmed rColXVII-C shows a much higher melting temperature as compared with the recombinant C15 domain expressed without the portion of the NC16a domain (21). In this light it is interesting that a stretch of three heptad repeats preceding the Gly-X-Y repeats show a probability to form a coiled-coil structure when analyzed by the COILS program (26). It has previously been hypothesized that 39 amino acid residues (positions 471–509) close to the transmembrane region can form a three-stranded coiled-coil that may play a role in assembly (7). Analysis of this region with the COILS program results in a very high probability to form coiled-coils. However, in fragment rColXVII-C, this region is not included. It may be possible that *in vivo* both regions may play a role in aligning the α-chains in a proper stagger.

Triple helix formation of fibril-forming collagens starts at the C-terminal end of the molecule (8, 27). Collagen XVII is a type II transmembrane protein with an intracellular N and an extracellular C terminus. It has been suggested that unlike other collagens, the membrane-bound collagens fold from the N- to the C-terminal end (28). The data presented here support this hypothesis for several reasons as follows. (i) The deletion of the C-terminal NC1 domain does not abolish triple-helix formation of rColXVII-B. (ii) The small C-terminal fragment rColXVII-D, without the context of more N-terminal sequences, does not adopt a triple-helical structure despite the presence of several collagenous stretches. (iii) The N-terminal C15 collagenous domain can be recombinantly expressed as a triple-helical polypeptide, although it lacks the entire C-terminal portion of the ectodomain. (iv) The stability of this polypeptide increases when N-terminal amino acid residues from the NC16a domain are present.

Acknowledgments—We thank Dr. L. Bruckner-Tuderman and Dr. T. Tasanen for the generous gift of antibodies to the NC16a domain of collagen XVII, Dr. E. Kohfeldt for the pCEP-Pu expression vector, and Dr. G. J. Giudice for the collagen XVII cDNA. We acknowledge the excellent technical assistance of G. Baines and Karin Wieβmann. M. P. Marinkovich acknowledges the Office of Research and Development, Veterans Affairs Palo Alto Health Care System, and the Dermatology Foundation.

REFERENCES

1. Jones, J. C., Hopkinson, S. B., and Goldfinger, L. E. (1998) *Bioessays* **20,** 488–494
2. Borradori, L., and Sonnenberg, A. (1999) *J. Invest. Dermatol.* **112,** 411–418
3. Bruckner-Tuderman, L. (2000) in *Connective Tissue and Its Heritable Disorders: Molecular, Genetic, and Medical Aspects* (P. Royce, and B. Steinmann, eds) John Wiley & Sons, Inc., New York, in press
4. Giudice, G. J., Emery, D. J., and Diaz, L. A. (1992) *J. Invest. Dermatol.* **99,** 243–250
5. Schäcke, H., Schumann, H., Hammami-Hauasli, N., Raghunath, M., and Bruckner-Tuderman, L. (1998) *J. Biol. Chem.* **273,** 25937–25943
6. Hirako, Y., Usukura, J., Uematsu, J., Hashimoto, T., Kitajima, Y., and Owaribe, K. (1998) *J. Biol. Chem.* **273,** 9711–9717
7. Balding, S. D., Diaz, L. A., and Giudice, G. J. (1997) *Biochemistry* **36,** 8821–8830
8. Olsen, B. R., Hoffmann, H., and Prockop, D. J. (1976) *Arch. Biochem. Biophys.* **175,** 341–350
9. Bulleid, N. J., Dalley, J. A., and Lees, J. F. (1997) *EMBO J.* **16,** 6694–6701
10. Lees, J. F., Tasab, M., and Bulleid, N. J. (1997) *EMBO J.* **16,** 908–916
11. McLaughlin, S. H., and Bulleid, N. J. (1998) *Matrix Biol.* **16,** 369–377
12. Mazzorana, M., Giry-Lozinguez, C., and van der Rest, M. (1995) *Matrix Biol.* **14,** 583–588
13. Kohfeldt, E., Maurer, P., Vannahme, C., and Timpl, R. (1997) *FEBS Lett.* **414,** 557–561
14. Chen, C., and Okayama, H. (1987) *Mol. Cell. Biol.* **7,** 2745–2752
15. Laemmli, U. K. (1970) *Nature* **227,** 680–685
16. Giudice, G. J., Emery, D. J., Zelickson, B. D., Anhalt, G. J., Liu, Z., and Diaz, L. A. (1993) *J. Immunol.* **151,** 5742–5750
17. Schumann, H., Baetge, J., Tasanen, K., Wojnarowska, F., Schäcke, H., Zillikens, D., and Bruckner-Tuderman, L. (2000) *Am. J. Pathol.* **156,** 685–695
18. Fietzek, P. P., Rexrodt, F. W., Wendt, P., Stark, M., and Kühn, K. (1972) *Eur. J. Biochem.* **30,** 163–168
19. Burgeson, R. E. (1993) *J. Invest. Dermatol.* **101,** 252–255

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

20. Bruckner, P., and Prockop, D. J. (1981) *Anal. Biochem.* **110,** 360–368
21. Tasanen, K., Eble, J. A., Aumailley, M., Schumann, H., Baetge, J., Tu, H., Bruckner, P., and Bruckner-Tuderman, L. (2000) *J. Biol. Chem.* **275,** 3093–3099
22. Hirako, Y., Usukura, J., Nishizawa, Y., and Owaribe, K. (1996) *J. Biol. Chem.* **271,** 13739–13745
23. Goldfinger, L. E., Hopkinson, S. B., deHart, G. W., Collawn, S., Couchman, J. R., and Jones, J. C. (1999) *J. Cell Sci.* **112,** 2615–2629
24. Hopkinson, S. B., Riddelle, K. S., and Jones, J. C. (1992) *J. Invest. Dermatol.* **99,** 264–270
25. Borradori, L., Koch, P. J., Niessen, C. M., Erkeland, S., van Leusden, M. R., and Sonnenberg, A. (1997) *J. Cell Biol.* **136,** 1333–1347
26. Lupas, A., Van Dyke, M., and Stock, J. (1991) *Science* **252,** 1162–1164
27. Beck, K., and Brodsky, B. (1998) *J. Struct. Biol.* **122,** 17–29
28. Pihlajaniemi, T., and Rehn, M. (1995) *Prog. Nucleic Acid Res. Mol. Biol.* **50,** 225–262

Downloaded from http://www.jbc.org/ by guest on January 24, 2014

View publication stats