J. Biochem. 128, 957–963 (2000)

Production in Mammalian Cells of Chimeric Human/Sea Urchin Procollagen Molecules Displaying Distinct Versions of the Minor Triple Helix[1]

Caroline Cluzel, Claire Lethias, Robert Garrone, and Jean-Yves Exposito[2]

Institut de Biologie et Chimie des Protéines, CNRS, Unité Mixte de Recherche 5086, Université Claude Bernard, 7 passage du Vercors, 69367 Lyon cedex 07, France

Received July 24, 2000; accepted September 22, 2000

One of the mechanisms involved in the regulation of the fibril diameter is the retention of the *N*-propeptide. In sea urchin embryo, thin collagen fibrils harbor numerous extensions at their surface, which we have suggested correspond to the large *N*-propeptide of the 2α collagen chain. To investigate the function of the *N*-propeptide during fibrillogenesis, we engineered constructs coding for the globular region of the 2α *N*-propeptide. To obtain homotrimeric molecules, the *N*-telopeptide, the central triple helix and the *C*-propeptide of the 2α chain were replaced by human domains of the proα1(I) chain. A single restriction site allowed insertion of distinct versions of the minor triple helix of the *N*-propeptide. Several human cell lines were transfected, and with one of them we were able to produce intact homotrimeric procollagen molecules. Rotary shadowing of these purified molecules indicates the presence of three large 2α *N*-propeptides that are similar to the extensions present at the surface of the sea urchin thin fibrils. This cassette-vector will be useful in determining the respective contributions of the globular and minor triple helical domains of the *N*-propeptide in the regulation of fibril diameter.

Key words: amino-propeptide, collagen, fibril, homotrimer, recombinant protein.

Collagen is one of the major components of the extracellular matrix of all multicellular organisms (*1*). All collagen molecules are made of three α chains, identical or not, containing at least one domain made of successive Gly-Xaa-Yaa triplets. These three chains are coiled around one another in a triple helical conformation. Among the collagens, the major family consists of the so-called fibrillar collagens, including five types in vertebrates (types I–III, V, and XI). All of the corresponding α chains share the same overall structure, consisting of a large uninterrupted central triple helix of approximately 338 Gly-Xaa-Yaa triplets (the main triple helix), flanked by two non-collagenous extensions, the *N*- and *C*-propeptides. During extracellular maturation of procollagens into collagen molecules, the *N*- and *C*-propeptides are generally removed by the action of specific proteases. The last step is the formation of the quarter-staggered fibrils (*2*, *3*). Fibril shape varies among the tissues and/or the collagen types involved (*4*). In tendon, the fibrils show a large spectrum of diameters and are involved in the mechanical strength of this tissue. In cornea, fibrils are organized into orthogonal layers and show a uniformly thin diameter. Their association with proteoglycans seemed to be necessary for the cornea transparency.

Several mechanisms have been proposed or defined which could govern fibril diameter. One is the presence of proteoglycans like decorin. When *in vitro* fibrillogenesis is carried out in the presence of decorin, the resulting fibrils are thinner (*5*), and in decorin-deficient mice, they display an irregular aspect (*6*). When pepsinized collagen molecules are used in *in vitro* fibrillogenesis, the average diameter is distinct for each type of collagen, indicating an intrinsic control of diameter by the primary structure and/or glycosylation of the central triple helix (*7*, *8*). Absence of α2(I) chains in *oim*/*oim* (osteogenesis imperfecta-murine) mice results in a loss of the normal lateral packing in collagen fibrils in tail tendon (*9*). In heterotypic type I/V fibrils, fibril diameter depends on the ratio between these two collagen types, where diameter decreases as the percentage of type V molecules increases. In these fibrils, the maintenance of type V amino-propeptide seems to be the element responsible for the regulation of fibril diameter (*10*). In Linsenmayer's model, the *N*-telopeptide acts as a flexible hinge, which allows the projection of the globular part of the *N*-propeptide onto the fibril surface. In this model, the short triple helix extends to the hole zone resulting from the quarter-staggered arrangement of collagen molecules. Recently, Gregory *et al.* (*11*) made a structural analysis of the type XI *N*-propeptide and presented a model related to Linsenmayer's. These authors suggested that the globular region of the α1(XI) *N*-propeptide is projected outside the fibril either by the p6a/p6b-p8 region or by the minor triple helix. Therefore, distinct mechanisms appear to govern collagen fibril diameter depending on the tissue analyzed and the collagen types involved.

Rotary shadowing experiments using isolated fibrils from sea urchin plutei revealed that they share a thin diameter

[1] This work was supported in part by the CEE contract "Biotechnology B104-CT96-OG62".Caroline Cluzel is supported by the Fondation Marcel Mérieux and the Ligue Contre le Cancer-Comité de la Drôme.

[2] To whom correspondence should be addressed. Tel: +33-4-72-72-26-74, Fax: +33-4-72-72-26-02, E-mail: jy.exposito @ibcp.fr

and harbor numerous extensions at their surface (*12*). These extensions consist of a rod-like structure terminated by a small globule. From these data and the ultrastructural immunolabeling results (*12*), we suggested that these extensions represent the large non-collagenous region of the *N*-propeptide of the 2α chain, and that like for type V collagen, the maintenance of the *N*-propeptide could be correlated with the formation of thinner fibrils. The goal of this study was to engineer a cassette-construct permitting the production in mammalian cells of fibrillar collagen chains harboring distinct versions of the *N*-propeptide. The production of these homotrimeric recombinant molecules is a prerequisite step in the analysis of the respective contributions to fibril diameter of the globular and the minor triple helix of the *N*-propeptide.

MATERIALS AND METHODS

Recombinant Protein and Polyclonal Antibody Production—We used a derivative of the plasmid pT7/7 (United States Biochemical), in which 6 His codons have been included between the *Pst*I and *Hin*dIII sites (*13*). The DNA insert, encoding the SURF modules R1–R5 and R9–R10 from the sea urchin 2α fibrillar collagen chain, was generated by PCR using the cDNA clone F6 as template (*14*) with Goldstar DNA polymerase in standard conditions (Eurogentec, Seraing). The oligonucleotides included an *Eco*RI site in the 5′ primer (5′TATGAATTCCGGTGATCACCAGTACAGGCA3′) and a *Pst*I site in the 3′ primer (5′TATCTGCAGTCTTCGGCAAGGCTGCTCCTT3′). The *Eco*RI–*Pst*I digested DNA was ligated into the expression vector digested with the same enzymes. For protein production in the *Escherichia coli* strain BL21(DE3) and purification of the 6-His–tagged recombinant protein, we used previously described conditions (*12*). Polyclonal antibodies (called BII) were raised in guinea pigs at the Valbex Center (Institut Universitaire de Technologie, Villeurbanne, France) using standard immunisation procedures.

DNA Constructs Encoding Chimeric Sea Urchin/Human Collagen Chains—The cDNA 1α3 and clone 5 (*15*) encoding part of the human proα1(I) collagen chain and cDNAs F6 and f3 (*14*) encoding part of the sea urchin 2α chain were used. The 3′ end of the construct consisted of a *Xba*I–*Hin*dIII fragment modified from the clone 5. The *Xba*I–*Dra*III fragment was replaced by the *Xba*I/*Pst*I–*Dra*III PCR product amplified using 1α3 as template and primers designed to introduce a *Pst*I restriction site at the 5′ end and to generate a mutated aminoproteinase cleavage site (Pro-Gln→Pro-Gly) (sense primer, 5′TATTCTAGACTGCAGCTCCCGGTCTGTCTTATGGCTATGATGAGAAATCAACCGGAGGAATTTCCG3′; anti-sense primer, 5′TGAGGTTGTAGAAGTTCCG3′). This modified *Xba*I–*Hin*dIII fragment (clone A) encoded the *N*-telopeptide, triple helical and *C*-propeptide domains of the human proα1(I) chain. To obtain the 5′ end of the construct, the *Eco*RI/*Xba*I/*Hin*dII–*Kpn*I PCR fragment (sense primer, 5′TATGAATTCTAGAAGCTTAGACATGTTCAGCTTTGTGGAC3′; anti-sense primer, 5′GTTTCCACACGTCTCGGTCATGGT3′) was cloned into pUC19 (clone B). This *Eco*RI–*Kpn*I fragment encoded the translation initiation codon, the signal peptide and part of the Tsp-2 domain of the human proα1(I) chain. The *Kpn*I–*Pst*I fragment of F6 cDNA encoding the SURF domains R1 to R5 was ligated into the *Kpn*I and *Pst*I sites of clone B in order to obtain clone C. The sequence coding for the Tsp-2 domain was ligated into the *Kpn*I site of clone C with a *Kpn*I-*Kpn*I PCR product amplified from the F6 cDNA (sense primer, 5′AACGGTACCTCCACGCAGAATCGTGGAGTG3′, anti-sense primer, 5′CCGTTGTATCCTGAGCTTCCTGGA3′), generating clone D. The chimeric TH0 construct encoding the αTH0 chain was obtained by ligating the *Xba*I–*Pst*I fragment of clone D into the same sites of clone A. This construct encodes a hybrid procollagen chain without the short triple-helical domain of the *N*-propeptide. The *Pst*I–*Pst*I PCR cassette encoding the short triple-helical domain of the 2α *N*-propeptide was ligated into the *Pst*I site of TH0 (sense primer, 5′TATCTGCAGAATGTGCCAAGGCTGCTCCACCTCCAGTGCAAG3′; anti-sense primer, 5′TATCTGCAGCAGTCATCTCACCGGGAG3′). The resultant construct, called TH1, encodes the αTH1 hybrid procollagen chain. To generate the TH2 construct, two further steps were necessary. The *Pst*I–*Bam*HI PCR product amplified from f3 cDNA (sense primer, 5′TATCTGCAGAATGTGCCAAGGCTGCTCCACCTCCAGTGCAAG3′; anti-sense primer, 5′TATGGATCCCATCTCACCGGGAGGGCCAGGAGGTCCAG3′) was ligated into the same sites of pUC19, generating clone E, which encodes an interrupted triple helical domain of the *N*-propeptide with the triplet sequence [(Gly-Xaa-Yaa)$_4$-5aa-(Gly-Xaa-Yaa)$_{20}$]. A second PCR product, *Hin*dIII/*Bam*HI–*Pst*I (sense primer, 5′TATAAGCTTGGATCCGTCGGTATGGTTGGACCCCCTGGAAGAC3′; anti-sense primer, 5′TATCTGCAGCAGTCATCTCACCGGGAG3′), was cloned into *Hin*dIII–*Pst*I sites of pUC19 and then excised by *Bam*HI, using the site present in the pUC19 polylinker. This *Bam*HI–*Bam*HI fragment was inserted into the same site of clone E, adding the sequence coding for 21 further triplets at the C-terminal end of the short triple helix. The resultant clone F was digested by *Pst*I and the *Pst*I–*Bam*HI–*Pst*I cassette was ligated into the unique site of TH0. The resultant construct TH2 encodes the αTH2 chain, which contains a twice normal size triple helical domain in its *N*-propeptide. The triplet sequence of this domain is [(Gly-Xaa-Yaa)$_4$-5aa-(Gly-Xaa-Yaa)$_{41}$], where the 41 triplets consist of two stretches of 20 triplets of the 2α *N*-propeptide triple helix with one central triplet Gly-Ser-Pro corresponding to the *Bam*HI junction of the two PCR fragments.

The full hybrid constructs were excised with *Hin*dIII. Blunt-ended fragments were generated with Klenow and inserted into the *Eco*RV site of the mammalian bicistronic expression vector pIRES1neo (Clontech), generating pIRES-αTH0, pIRES-αTH1, and pIRES-αTH2. The PCR products and cloning junctions were verified by DNA sequencing ("T7 sequencing mixes" Pharmacia).

Cell Culture and Transfection—Cells were cultured in Dulbecco's Modified Eagle's Medium from Sigma (DMEM) supplemented with 10% (v/v) fetal calf serum (Seromed) and penicillin/streptomycin from Sigma (100 U/100 μg/ml). Human fibrosarcoma HT1080 cells (ATCC CCL121) were transfected with DNA constructs using Transfectam Reagent (Promega). Briefly, cells were grown to confluency in a 60-mm cell culture dish and provided with 500 μl of medium without fetal calf serum. The cells were incubated for 2 h in the presence of 1 ml of serum-free DMEM with 2.6 μg of DNA and 21 μl of Transfectam. After incubation, 1.5 ml of medium with 20% fetal calf serum was added. Human amnion WISH cells (ATCC CCL25) were trans-

fected by electroporation (500 μF, 400 V) with 60 μg of DNA for a confluent 100-mm cell culture dish, then seeded in a 140-mm cell culture dish. Human embryonic kidney 293 cells (HEK293, ATCC 1573) were transfected by phosphate-calcium as described (*16*).

Screening of Transfected Clones—After 48 h, transfected cells were selected by the addition of G418 at a concentration of 1 mg/ml and incubation for about 2 weeks. Isolated G418-resistant clones were expanded in 96- and then 24-well plates. For screening, 1 ml of medium from each confluent well was collected and proteins were precipitated with 5% polyethylene glycol 8000 (*17*). After centrifugation, pellets were resuspended in sample electrophoresis buffer. Proteins were separated by 6% SDS-PAGE followed by electroblotting onto PVDF. The presence of recombinant procollagen was assayed by Western blot analysis with polyclonal guinea pig anti-SURF antibodies (BII) or rabbit anti–bovine collagen I antibodies (Ib) and secondary antibodies of peroxidase-conjugated sheep anti–guinea pig (Biosys) or alkaline phosphatase–conjugated goat anti-rabbit (Dako). Peroxidase was revealed by luminescence using the Renaissance kit (NEN Life Science Products) and alkaline phosphatase was detected by the NBT/BCIP method using the Alkaline Phosphatase Conjugate Substrate Kit (Biorad).

Recombinant Procollagen Production and Purification—Recombinant 293 cell clones secreting chimeric procollagen were grown in DMEM supplemented with 5% serum in 100-mm cell culture dishes. Confluent cells were washed with phosphate-buffered saline (PBS), and 5 ml of serum-free medium containing 50 μg/ml ascorbate was added and collected every 24 h. The purification of recombinant collagen was achieved according to the method of Zafarullah *et al.* (*18*) with some modifications. The medium was supplemented with 1:10 (v/v) stock solution (1 M Tris-HCl pH 7.4, 250 mM EDTA, 0.2% NaN_3, 0.1 mM DTT) and precipitated overnight by ammonium sulfate (176 mg/ml). After centrifugation at 20,000 ×*g* for 1 h, the pellet was re-suspended in storage buffer (0.1 M Tris-HCl pH 7.4, 0.4 M NaCl, 0.04% NaN_3, 0.1 mM DTT) and stirred overnight. The suspension was centrifuged (20,000 ×*g*, 15 min) and dialyzed against buffer I (0.1 M Tris-HCl pH 7.5, 2 M Urea,

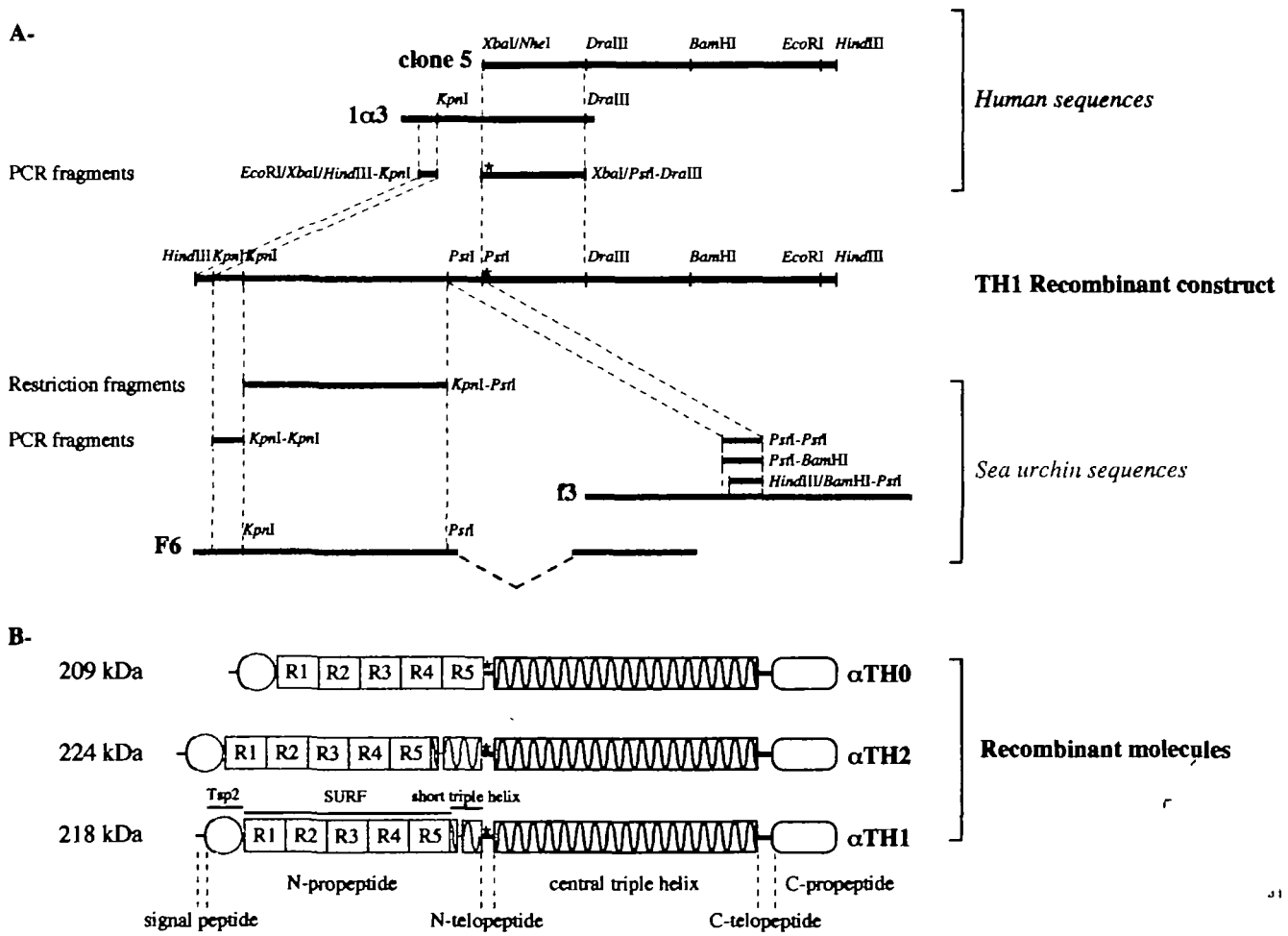

Fig. 1. **Chimeric procollagen constructs.** A: Restriction fragments and PCR products used to generate TH0, TH1, and TH2 expression constructs are represented below the human DNA templates (black lines) or above the sea urchin cDNAs (grey lines). In the middle, the resultant chimeric expression construct TH1 is shown. B: Composition of the engineered recombinant molecules. From the native α chain (αTH1), two mutated chains are shown. For αTH0 the short triple helix of the *N*-propeptide is absent, while this domain has doubled in size for the αTH2 chain. The theoretical molecular mass of the chains is indicated on the left side; the mutated aminoproteinase site is indicated by an asterisk (*).

0.2 M NaCl, 0.04% NaN_3, 0.1 mM DTT), then the sample was loaded onto a DEAE-cellulose anion-exchanger (DE52, Whatman) equilibrated with buffer I. The flow-through fraction was dialyzed against buffer II (0.15 M NaCl, 25 mM EDTA, 0.02% NaN_3, 0.05 M Tris-HCl pH 7.0, 0.1 mM DTT) and passed through a gelatin-Sepharose column in order to remove fibronectin (Pharmacia). The flow-through fraction was dialyzed against buffer III (0.075 M Tris-HCl pH 7.8, 2 M urea, 0.04% NaN_3, 0.1 mM DTT) and loaded onto a second column of DEAE-cellulose (DE52), which was eluted with a linear gradient of 0 to 0.2 M NaCl. Fractions containing recombinant procollagen were pooled and dialyzed against storage buffer and then concentrated by ultrafiltration with a 100 kDa molecular mass cut-off membrane (Ultrafree, Millipore).

Proteolytic Digestion—To assay the protease resistance of the recombinant procollagen, 1 ml of medium was preincubated at 37–45°C for 30 min and digested at the same temperature for 10 min with trypsin using an enzyme/substrate ratio of 1/100. Triple-helical resistant chains were analyzed by Western blotting using Ib antibodies.

Rotary Shadowing—The procollagen solution was mixed with glycerol (v/v) and spread on freshly cleaved mica sheets. The sample was shadowed in a Balzers MED10 coating unit. Platinum-carbon replicas were observed with a Philips CM120 microscope at the Centre de Microscopie Electronique Appliquée à la Biologie et à la Géologie, Université Claude Bernard, Lyon.

RESULTS

In plutei, collagen fibrils have a uniform diameter of 25 nm and large extensions are observed periodically at their surface. We suggested previously that these extensions correspond to the 2α *N*-propeptide (*12*) and that, in the same way as in vertebrate types V and XI, the steric hindrance of this domain limits the growth of the fibril.

The aim of this study was to produce recombinant collagen harboring the 2α *N*-propeptide in a system where the *N*-propeptide was not cleaved and where different configurations of the minor triple helix of this domain could be easily generated. In this way, the functions of the globular and minor triple helical regions of the *N*-propeptide in the regulation of the fibril diameter could be investigated. To achieve this goal, the construct should possess a unique restriction site in which the sequence encoding the minor triple helix could be included, as well as a mutation in the sequence encoding the *N*-proteinase site. Since the 2α chain is involved in the formation of heterotrimeric structure [$(1\alpha)_2 2\alpha$] (*19*), we replaced the *N*-telopeptide, the main triple helix and the *C*-propeptide coding sequences of the 2α chain by its human proα1(I) counterpart. All cloning steps and the general structure of the recombinant chimeric α chains are represented in Fig. 1. For the control chain αTH1, the *N*-propeptide has the native structure with a Tsp-2 module, five SURF domains and the minor triple helix [$(Gly\text{-}Xaa\text{-}Yaa)_4$-5aa-$(Gly\text{-}Xaa\text{-}Yaa)_{20}$] of the sea urchin 2α chain. Two mutant isoforms were designed, differing in the minor triple helix of the *N*-propeptide. This segment is absent in αTH0, whereas for the αTH2 chain, it is twice its normal size [$(Gly\text{-}Xaa\text{-}Yaa)_4$-5aa-$(Gly\text{-}Xaa\text{-}Yaa)_{41}$]. Three human cell lines were tested, and the results obtained with αTH1 are shown in Fig. 2. For its production, pIRES-αTH1 was first transfected in HT1080 cells, which have been previously used with success in the production of type I and type II recombinant collagens (*17*, *20*). As shown in Fig. 2A, only small amounts of intact recombinant proteins were obtained using this cell line. Using anti-type I antibodies, the major product had an apparent molecular mass of 80 kDa. Following these results, we used WISH cells, which are known to produce type VII and type XII collagens (*21*). As presented in Fig. 2A, the *N*-propeptide and probably the *C*-propeptide were fully processed in this recombinant assay, even though the *N*-proteinase site had been mutated. Using anti-type I antibodies (Fig. 2A), the major positive product co-migrated with the α1 chain of a bovine type I standard. A positive band of 75 kDa was detected using anti-2α *N*-propeptide antibodies (Fig. 2A). Its apparent mobility agreed closely with the theoretical molecular mass of the 2α *N*-propeptide of the αTH1 chain. In a third step, we used HEK293 cells, which had been used for the recombinant production of non-fibrillar collagens (*22–24*) and fibrillar collagens (*25*, *26*) and appeared to have prolyl-4-hydroxylase activity (*27*). In Fig. 2A, the major product had an apparent molecular mass in agreement with the theoretical value (218 kDa) of αTH1. A weaker and faster migrating band of 215 kDa was also observed. The production of not only αTH1 but also αTH2

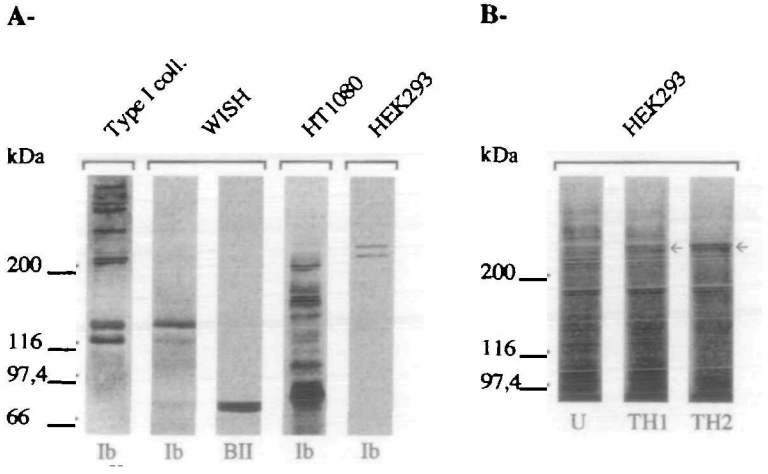

Fig. 2. **Analysis of culture medium from different human cell lines transfected by TH1-pIRES.** A: Western blot analysis of αTH1 secreted by transfected HT1080, WISH, and HEK293 cells. Recombinant proteins were precipitated from the media, separated by 6% SDS-PAGE and electrotransferred. They were then incubated with polyclonal rabbit anti-bovine collagen I antibodies (Ib) or polyclonal guinea pig anti-SURF antibodies (BII) and detected with an alkaline phosphatase (Ib) or peroxidase (BII) conjugate. For comparison, a sample of purified bovine heterotrimeric collagen I (Type I coll.) was also loaded. B: SDS-PAGE of medium from HEK293 cells, untransfected (U) or transfected by TH1-pIRES (TH1) or TH2-pIRES (TH2). Proteins from 1 ml of medium were analyzed by 5% SDS-PAGE and stained with Coomassie Blue.

could be detected in Fig. 2B, where crude extracts were analyzed by SDS-PAGE and compared to untransfected cell extracts. In Fig. 3A, Western-blot analysis of the αTH0, αTH1, and αTH2 chains using anti-2α *N*-propeptide antibodies was performed in reducing or non-reducing conditions. Without reduction, the slowest migrating product of each recombinant chain was converted into the band of highest molecular mass due to the maintenance of inter-chain disulfide bonds. For the αTH2 chain, in some cases an additional product was observed which was not converted into a slower migrating band under non-reducing conditions (see Fig. 3A). This product was rarely observed and probably corresponds to the αTH2 chain without its *C*-propeptide, consistent with the molecular mass difference between the two forms. From these results, the inter-chain disulfide bonds appeared to occur in the *C*-propeptides. The thermal stability of the αTH1 molecules was monitored by trypsin digestion and a T_m of 41–42°C was observed, indicating a correct triple helical conformation of the molecules (Fig. 3B). Similar values were obtained with αTH0 and αTH2 molecules (data not shown). Purification of αTH1 and αTH2 recombinant molecules was performed, and the purified products were analyzed by rotary shadowing. Electron microscopic observation revealed molecules with a thread-like structure of 300 nm on average, flanked by a globule which corresponds to the *C*-propeptide, and three extensions, each of which correspond to an *N*-propeptide (Fig. 4A). As indicated in the legend of Fig. 4, the triple helix region of αTH2 molecules (314 ± 6 nm) is longer than

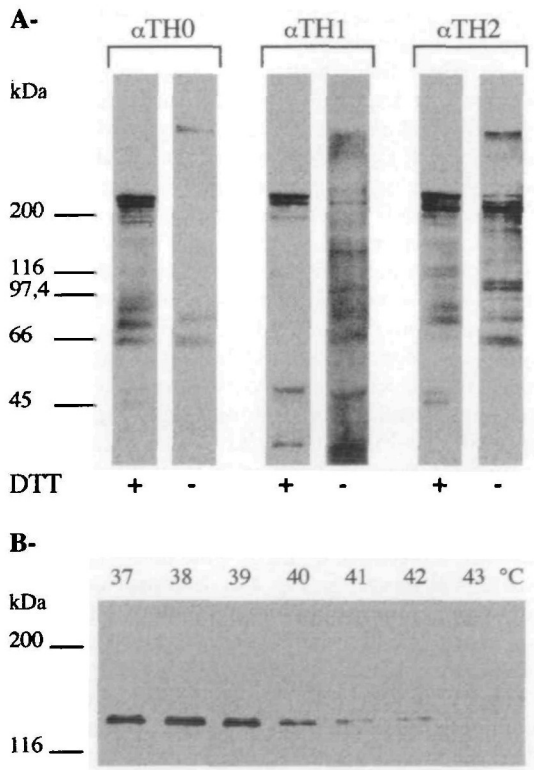

Fig. 3. **Analysis of chimeric recombinant chains produced by HEK293 cells.** A: Western-blot analysis with BII antibodies of αTH0, αTH1, and αTH2 chains with (+DTT) or without (–DTT) reduction. B: Thermal stability of homotrimeric αTH1 molecules. Medium from transfected cells was digested with trypsin at the indicated temperature. Resistant molecules were separated by SDS-PAGE, transferred to the membrane and reacted with Ib antibodies. DTT, dithiothreitol.

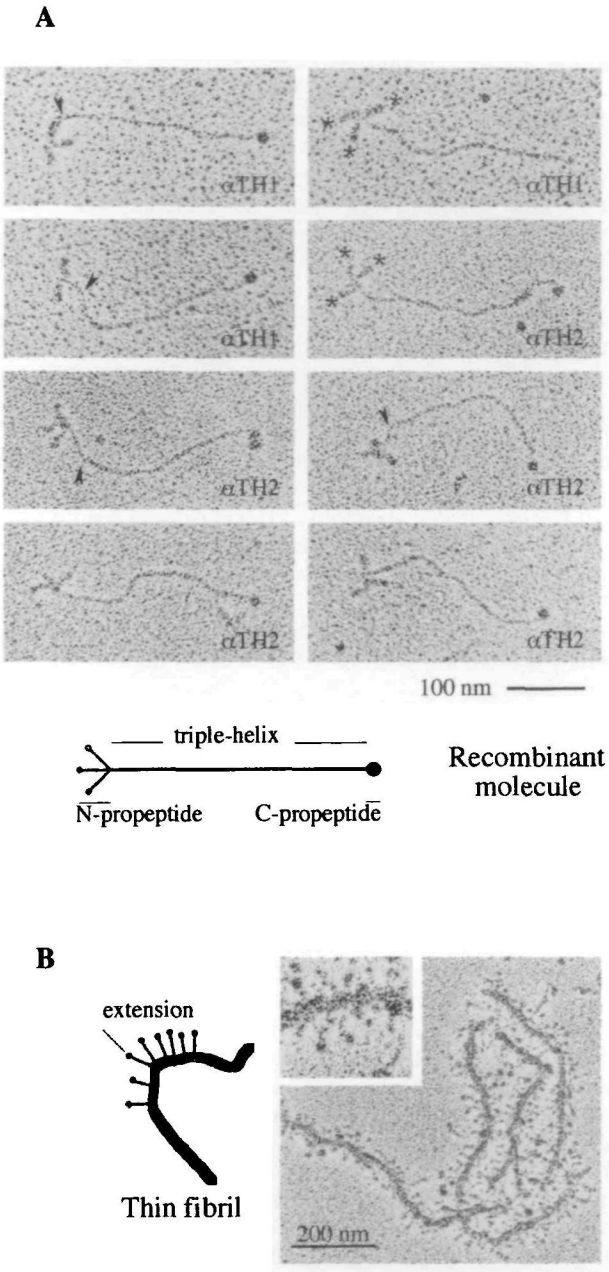

Fig. 4. **Electron microscopy of recombinant collagen molecules and sea urchin fibril.** A: αTH1 and αTH2 molecules visualized after rotary shadowing by electron microscopy. The mean lengths of the triple helix region of the molecules are: αTH1 = 296 ± 8.4 nm (n = 17) and αTH2 = 314 ± 6.5 nm (n = 27). Arrowheads point out possible bends which might reflect a flexible hinge region between the minor and the major triple helix. Typical rod-like structures terminated by a globule are marked by asterisks. A diagrammatic representation of a recombinant collagen molecule is presented below the electron microscopy pictures. B: Thin sea urchin fibril vizualized after rotary shadowing by electron microscopy. Note the presence of periodic extensions at the surface of this fibril which present the same structural features as the 2α *N*-propeptides of the αTH1 and αTH2 molecules shown in A. Insert: magnification (twice) of a region of the fibril. A schematic representation of a thin fibril is shown at the left of the rotary shadowing image.

that of αTH1 molecules (296 ± 8.4 nm). The difference in size between the two recombinant molecules should correspond to the increased number of Gly-Xaa-Yaa triplets in the minor triple helix of αTH2 compared to αTH1. For the two molecular types, the triple helix structure shows several bends occasionally occuring in a region which might be located at the junction between the minor and the major triple helix (Fig. 4A). The *N*-propeptide extensions contained a rod-like structure corresponding to the five SURF module repeats. The rod-like structure ended with a small globule, which most likely corresponded to the Tsp-2 domain. This structure was similar to the extensions periodically projecting from thin sea urchin fibrils (Fig. 4B) (*12*). From these experiments and the previous immunolocalization of the 2α *N*-propeptide at the surface of thin fibrils (*12*), we have indirect confirmation that they should correspond to the *N*-propeptide of the 2α chain.

DISCUSSION

In this report, electron microscopic analysis of intact chimeric recombinant molecules has allowed us to conclude that the extensions observed at the surface of sea urchin 25 nm diameter fibrils are the non-collagenous domains of the 2α *N*-propeptide. Construction of a cassette-vector and use of the HEK293 cell line will permit the analysis of the respective contributions of the non-collagenous and triple helix parts of the *N*-propeptide in the regulation of fibril diameter.

Rotary shadowing pictures of homotrimeric recombinant molecules strongly support our previous suggestion that the extensions protruding from thin sea urchin fibrils correspond to the globular region of the 2α *N*-propeptide (*12*). It should be noted that only one extension is present in these fibrils and not three as in recombinant molecules. This is in agreement with the molecular composition (*19*) identified in sea urchin [$(1\alpha)_2 2\alpha$]. Hence, we have two 1α chains (*28*) with an *N*-propeptide lacking a globular region, and one 2α chain possessing a large globular domain in its *N*-propeptide.

Of the different human cell lines used to produce the chimeric recombinant α chains, only one was suitable for obtaining intact molecules. Hence, mutation of the *N*-proteinase site is not sufficient to prevent the processing of the *N*-propeptide in WISH cells. We believe that the region including the minor triple helix is not stably folded and is more accessible to specific cleavage by proteases. As shown in Fig. 2, only the HEK293 cell line gives satisfactory results in terms of quality, since the molecules are intact, but also in terms of quantity, since the recombinant molecules could be visualized by Coomassie Blue staining in crude extracts. The major problem inherent in this cell line is the persistence of the *C*-propeptide as shown by non-reducing SDS-PAGE analysis and by rotary shadowing. In previous reports, the *C*-propeptide of homotrimeric recombinant proα1(V) collagen was rapidly processed in 293-EBNA cell culture (*25*, *26*). Imamura *et al.* (*26*) suggest that a furin-like proprotein convertase processes this domain. The human proα1(I) chain does not have a consensus site for this protease, yet the *C*-propeptide could be cleaved in some experiments, as shown for the αTH2 chain (Fig. 3). In fact, the processing of this domain is a prerequisite step for *in vitro* fibrillogenesis assays. From this result, the co-transfection of chimeric constructs with an expression vector carrying the sequence coding for BMP1, the types I and III *C*-proteinase (*29*, *30*), will permit us to obtain pNcollagens. The use of purified BMP1 enzyme and C-terminal proteinase enhancer could be an alternative approach to obtain correctly C-terminally processed molecules.

The principal interest of our cassette constructs is to provide the tools necessary to understand the function of the minor triple helix of the *N*-propeptide in the regulation of fibril diameter. As indicated in the introduction, several mechanisms have been proposed or defined which could govern fibril diameter. It is worth noting that the only domain conserved during evolution in the *N*-propeptide among all fibrillar collagen chains characterized is this minor triple helix (*12*, *31*). Moreover, in all collagen chains implicated in the formation of thin, constant-diameter fibrils, this domain is well conserved in size, structure and sequence. These chains include the vertebrate types II, V [including the newly characterized α3 chain (*32*)], and XI, and the sea urchin 1α and 2α chains. A consensus signature of this domain would be $(\text{Gly-Xaa-Yaa})_{3\text{-}5}$, a short interruption followed by 17 to 20 Gly-Xaa-Yaa triplets. The first three to five triplets include the sequence Gly-Xaa-Lys-Gly-Xaa-Lys-Gly-Glu-Pro. For these reasons we believe that the size, sequence and structure of this domain is responsible for the diameter observed of thin fibrils, and that the minor triple helix of the *N*-propeptide should also be defined as an intrinsic factor involved in the regulation of the fibril diameter. This hypothesis is one of the two suggestions proposed by Gregory *et al.* (*11*). In their model and in Linsenmayer's (*10*), the *N*-telopeptide region which links the minor and the major triple helix acts as flexible hinge. This hinge allows the minor triple helix passing through the hole zone to project the globular part of the *N*-propeptide at the surface of the fibril. Consistent with this hypothesis, a bend has been shown in rotary shadowing pictures of types V and XI procollagen molecules (*10*, *33*). However, no bend was observed in recombinant homotrimeric α1(V) molecules (*25*). Moreover, a structural study of the type XI *N*-propeptide (*11*) revealed that its non-collagenous region consists of a globular head and an extended tail of approximately 16 nm long. Hence, the bend observed in type XI molecules might correspond to the junction between the non-collagenous tail region and the minor triple helix of the *N*-propeptide. In Fig. 4A, we have also shown molecules which might possess a bend in a location possibly corresponding to the junction between the minor and the major triple helix. Nevertheless, from the general aspect of the recombinant molecules, we have some doubt about the relevance of this observation (Fig. 4A). In a previous study (*12*), we presented a molecule which is partially detached from a thin sea urchin fibril. In this case, the 2α *N*-propeptide extension is connected to the fibril by a thinner element, the triple helix, with no apparent bend. In fact, the flexibility of the amino-terminal region of the *N*-telopeptide has been only demonstrated for type I collagen (*34*). As shown in Fig. 4B, SURF modules of the 2α *N*-propeptide present a rod-like structure with an approximate length of 40 nm and a diameter of 4–5 nm as judged from the extensions present at the surface of thin sea urchin fibrils. Since the gap region has been defined to be 1.5/2.5 × 32 nm (*11*), SURF modules could not remain inside the fibrils. The location of the SURF modules at the surface of thin periodic fibrils

(*12*) corroborates this suggestion. Thus, only the minor triple helix could extend through the hole zone and expose the SURF modules at the surface of the fibril. The availability of this cassette system allows us to produce chimeric recombinant molecules harboring distinct forms of the minor triple helix and will be useful to test the function of this domain in the regulation of the fibril diameter by *in vitro* fibrillogenesis studies.

We gratefully acknowledge Emmanuelle Tillet and David J. Hulmes for the critical reading of the manuscript.

REFERENCES

1. Van der Rest, M. and Garrone, R. (1991) Collagen family of proteins. *FASEB J.* **5**, 2814–2823
2. Prockop, D.J. and Kivirikko, K.I. (1995) Collagens: molecular biology, diseases, and potentials for therapy. *Annu. Rev. Biochem.* **64**, 403–434
3. Kadler, K.E., Holmes, D.F., Trotter, J.A., and Chapman, J.A. (1996) Collagen fibril formation. *Biochem. J.* **316**, 1–11
4. Parry, D.A.D. and Craig, A.S. (1984) Growth and development of collagen fibrils in connective tissue in *Ultrastructure of the Connective Tissue Matrix* (Ruggeri, A. and Motta, P.M., eds.) pp. 34–64, Martinus Nighoff Publishers, Boston
5. Vogel, K.G. and Trotter, J.A. (1987) The effect of proteoglycans on the morphology of collagen fibrils formed *in vitro*. *Coll. Relat. Res.* **7**, 105–114
6. Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E., and Iozzo, R.V. (1997) Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J. Cell Biol.* **136**, 729–743
7. Birk, D.E. and Silver, F.H. (1984) Collagen fibrillogenesis *in vitro*: comparison of types I, II, and III. *Arch. Biochem. Biophys.* **235**, 178–185
8. Notbohm, H., Nokelainen, M., Myllyharju, J., Fietzek, P.P., Müller, P.K., and Kivirikko, K.I. (1999) Recombinant human type II collagens with low and high levels of hydroxylysine and its glycosylated forms show marked differences in fibrillogenesis *in vitro*. *J. Biol. Chem.* **274**, 8988–8992
9. McBride, D.J., Choe, V., Shapiro, J.R., and Brodsky, B. (1997) Altered collagen structure in mouse tail tendon lacking the α2(I) chain. *J. Mol. Biol.* **270**, 275–284
10. Linsenmayer, T.F., Gibney, E., Igoe, F., Gordon, M.K., Fitch, J.M., Fessler, L.I., and Birk, D.E. (1993) Type V collagen: molecular structure and fibrillar organization of the chicken α1(V) NH_2-terminal domain, a putative regulator of corneal fibrillogenesis. *J. Cell Biol.* **121**, 1181–1189
11. Gregory, K.E., Oxford, J.T., Chen, Y., Gambee, J.E., Gygi, S.P., Aebersold, R., Neame, P.J., Mechling, D.E., Bächinger, H.P., and Morris, N.P. (2000) Structural organization of distinct domains within the non-collagenous N-terminal region of collagen type XI. *J. Biol. Chem.* **275**, 11498–11506
12. Lethias, C., Exposito, J.Y., and Garrone, R. (1997) Collagen fibrillogenesis during sea urchin development: retention of SURF motifs from the *N*-propeptide of the 2α chain in mature fibrils. *Eur. J. Biochem.* **245**, 434–440
13. Cortay, J.C., Nègre, D., Scarabel, M., Ramseier, T.M., Vartak, N.B., Reizer, J., Saier, M., and Cozzone, A.J. (1994) *In vitro* asymmetric binding of the pleiotropic regulatory protein, FruR, to the ace operator controlling glyoxylate shunt enzyme synthesis. *J. Biol. Chem.* **269**, 14885–14891
14. Exposito, J.Y., D'Alessio, M., and Ramirez, F. (1992) Novel amino-propeptide configuration in a fibrillar procollagen undergoing alternative splicing. *J. Biol. Chem.* **267**, 17404–17408
15. Ruggiero, F., Exposito, J.Y., Bournat, P., Gruber, V., Perret, S., Comte, J., Olagnier, B., Garrone, R., and Theisen, M. (2000) Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Lett.* **469**, 132–136
16. Nischt, R., Pottgiesser, J., Krieg, T., Mayer, U., Aumailley, M., and Timpl, R. (1991) Recombinant expression and properties of the human calcium-binding extracellular matrix protein BM-40. *Eur. J. Biochem.* **200**, 529–536
17. Geddis, A.E. and Prockop, D.J. (1993) Expression of human *COL1A1* in stably transfected HT1080 cells: the production of a thermostable homotrimer of type I collagen in a recombinant system. *Matrix* **13**, 399–405
18. Zafarullah, K., Sieron, A.L., Fertala, A., Tromp, G., Kuivaniemi, H., and Prockop, D.J. (1997) A recombinant homotrimer of type I procollagen that lacks the central two D-periods. The thermal stability of the triple helix is decreased by 2 to 4°C. *Matrix Biol.* **16**, 245–253
19. Cluzel, C., Lethias, C., Garrone, R., and Exposito, J.Y. (2000) Sea urchin fibrillar collagen 2α chain participates in heterotrimeric molecules of $(1\alpha)_2 2\alpha$ stoechiometry. *Matrix Biol.* in press
20. Sieron, A.L., Fertala, A., Ala-Kokko, L., and Prockop, D.J. (1993) Deletion of a large domain in recombinant human procollagen II does not alter the thermal stability of the triple helix. *J. Biol. Chem* **268**, 21232–21237
21. Lunstrum, G.P., McDonough, A.M., Marinkovich, M.P., Keene, D.R., Morris, N.P., and Burgeson, R.E. (1992) Identification and partial purification of a large, variant form of type XII collagen. *J. Biol. Chem.* **267**, 20087–20092
22. Tillet, E., Wiedemann, H., Golbik, R., Pan, T.C., Zhang, R.Z., Mann, K., Chu, M.L., and Timpl, R. (1994) Recombinant expression and structural and binding properties of α1(VI) and α2(VI) chains of human collagen type VI. *Eur. J. Biochem.* **221**, 177–185
23. Tillet, E., Mann, K., Nischt, R., Pan, T.C., Chu, M.L., and Timpl, R. (1995) Recombinant analysis of human α1(XVI) collagen. *Eur. J. Biochem.* **228**, 160–168
24. Frischholz, S., Beier, F., Girkontaite, I., Wagner, K., Pöschl, E., Turnay, J., Mayer, U., and von der Mark, K. (1998) Characterization of human type X procollagen and its NC-1 domain expressed as recombinant proteins in HEK293 cells. *J. Biol. Chem.* **273**, 4547–4555
25. Fichard, A., Tillet, E., Delacoux, F., Garrone, R., and Ruggiero, F. (1997) Human recombinant α1(V) collagen chain. *J. Biol. Chem.* **272**, 30083–30087
26. Imamura, Y., Steiglitz, B.M., and Greenspan, D.S. (1998) Bone morphogenetic protein-1 processes the NH_2-terminal propeptide, and a furin-like proprotein convertase processes the COOH-terminal propeptide of pro-α1(V) collagen. *J. Biol. Chem.* **273**, 27511–27517
27. Hosokawa, N., Hohenadl, C., Satoh, M., Kühn, K., and Nagata, K. (1998) HSP47, a collagen-specific molecular chaperone, delays the secretion of type III procollagen transfected in human embryonic kidney cell line 293: a possible role for HSP47 in collagen modification. *J. Biochem.* **124**, 654–662
28. Exposito, J.Y., D'Alessio, M., Solursh, M., and Ramirez, F. (1992) Sea urchin homolog of Vertebrate pro-α2(I) collagen gene. *J. Biol. Chem.* **267**, 15559–15562
29. Li, S.W., Sieron, A.L., Fertala, A., Hojima, Y., Arnold, W.V., and Prockop, D.J. (1996) The *C*-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenetic protein-1. *Proc. Natl. Acad. Sci. USA* **93**, 5127–5130
30. Kessler, E., Takahara, K., Biniaminov, L., Brusel, M., and Greenspan, D.S. (1996) Bone morphognetic protein-1: the type I procollagen C-proteinase. *Science* **271**, 360–362
31. Exposito, J.Y., Cluzel, C., Lethias, C., and Garrone, R. (2000) Tracing the evolution of vertebrate fibrillar collagens from an ancestral α chain. *Matrix Biol.* **19**, 275–279
32. Imamura, Y., Scott, I.C., and Greenspan, D.S. (2000) The Pro-α3(V) collagen chain. Complete primary structure, expression domains in adult and developing tissues, and comparison to the structures and expression domains of the other types V and XI procollagen chains. *J. Biol. Chem.* **275**, 8749–8759
33. Thom, J.R. and Morris, N.P. (1991) Biosynthesis and proteolytic processing of type XI collagen in embryonic chick sterna. *J. Biol. Chem.* **266**, 7262–7269
34. Vitagliano, L., Némethy, G., Zagari, A., and Scheraga, H.A. (1995) Structure of the type I collagen molecule based on conformational energy computations: the triple-stranded helix and the N-terminal telopeptide. *J. Mol. Biol.* **247**, 69–80