Synthesis and conformational properties of a recombinant C-propeptide of human type III procollagen

ARTICLE *in* MATRIX BIOLOGY · NOVEMBER 1997

Impact Factor: 5.07 · DOI: 10.1016/S0945-053X(97)90009-3 · Source: PubMed

Regular Papers

MATRIX BIOLOGY

Synthesis and Conformational Properties of a Recombinant C-Propeptide of Human Type III Procollagen

KHAJA ZAFARULLAH, ELEANOR M. BROWN[1], HELENA KUIVANIEMI[2], GERARD TROMP[2], ALEKSANDER L. SIERON[3], ANDRZEJ FERTALA[3] and DARWIN J. PROCKOP[3]

Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, Jefferson Medical College, Thomas Jefferson University, Philadelphia, USA

Abstract

A cDNA was prepared that coded for the signal peptide of type III procollagen linked to the complete C-propeptide of the protein. The cDNA was then used to express the protein in a baculovirus recombinant system. Recombinant protein was recovered as a trimer from the medium of transfected cells in a yield of 1 to 2.5 mg per liter. Mapping of peptide fragments with and without reduction indicated that the protein contained the expected interchain disulfide bonds. Analysis by circular dichroism suggested that the conformation of the protein corresponded to the native conformation. Therefore, the protein should be appropriate for further tests of its biological function and analysis of structure by X-ray diffraction.

Key words: C-propeptide, baculovirus expression system, circular dichroism, collagen, extracellular matrix

Introduction

The most abundant collagens are the fibrillar types I, II and III (see Prockop and Kivirikko, 1995). They are similar in structure and function, and are also similar in that they are first synthesized as procollagens with N-terminal and C-terminal propeptide domains. In the processing of the procollagens to collagens, the C-propeptides are cleaved by a single procollagen C-proteinase (Hojima et al., 1985) that has recently been cloned and shown to have a structure identical to the protein previously defined as bone morphogenic protein-1 (BMP-1) (Kessler et al., 1996; Li et al. 1996). The N-propeptides of type I and type II procollagen are cleaved by a specific procollagen N-proteinase (Hojima et al., 1989) while the N-propeptide of type III procollagen is cleaved by a separate N-proteinase (Halila et al. 1986). After cleavage of the two propeptides, the resulting collagens spontaneously self-assemble into fibrils (see Prockop and Hulmes, 1994).

The C-propeptides of the fibrillar procollagens play a critical role in the biosynthesis of the proteins (Prockop

[1] Present address: United States Department of Agriculture, Eastern Regional Research Center, 600 E. Mermaid Lane, Philadelphia, PA.

[2] Present address: Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI.

[3] Present address: Center for Gene Therapy, Allegheny University of the Health Sciences, Philadelphia, PA.

Abbreviations used: BMP-1, bonemorphogenic protein-1; CD, circular dichroism; PABA, p-aminobenzamidine dihydrochloride acid; SFM, serum-free insect cell culture medium

and Kivirikko, 1995). They direct chain selection and stabilize the associated chains by the formation of interchain disulfide bonds. They also provide registration of the chains so that a nucleus of triple helix forms at the C-terminal end of the large collagen domain of the protein and the triple helix is then propagated from the C-terminus to the N-terminus (Schofield et al., 1974; Uitto and Prockop, 1974a,b; Bächinger et al., 1980). The critical role of the C-propeptides in type I procollagen in directing chain association was demonstrated both by assays *in vitro* and by naturally occurring mutations in man or mice that alter the structure of the C-propeptides and thereby prevent correct chain association (Nicholls et al., 1979; Deak et al., 1983; Chipman et al., 1993; Lamande et al., 1995; see also Kuivaniemi et al., 1997). In addition, there are suggestions that the C-propeptides, after cleavage from the procollagen molecules, serve as feed-back regulators of collagen biosynthesis either at the transcriptional level (Wu et al., 1986) or the post-transcriptional level (Aycock et al., 1986; Katayama et al., 1991; Wu et al., 1991). Also, there are indications that C-propeptides of type I procollagen bind to cell surfaces through the $\alpha 2\beta 1$ integrin and thereby play a role in cell attachment and spreading (Weston et al., 1994).

There is a high degree of amino acid homology among the C-propeptides of type I, type II and type III procollagens (Dion and Myers, 1987). Each of the propeptides has eight cysteine residues and an N-glycosylation site whose positions are highly conserved among the different chains. Also, the positions of the tyrosine, tryptophan and charged residues, as well as the patterns of hydrophilicity and hydrophobicity, are conserved (Dion and Myers , 1987). Studies *in vitro* on the C-propeptides of type I and type III procollagens suggested that cysteine residues at the amino terminal region form interchain disulfide bonds, whereas the more C-terminal cysteine residues are involved in the formation of intrachain disulfide bonds (Koivu, 1987; Lees and Bulleid, 1994).

Detailed structural analysis of the C-propeptide has been limited in part by the paucity of the protein available from natural sources. Here, we expressed a cDNA for the C-terminal propeptide of type III procollagen in a baculovirus expression system. The expressed recombinant protein was purified to homogeneity and assayed for formation of the correct disulfide bonds by cyanogen bromide mapping of the peptides under reducing and non-reducing conditions. Also, the conformation of the protein and its thermal stability were assayed by circular dichroism (CD).

Materials and Methods

Amplification and sub-cloning of type III procollagen C-propeptide

DNA constructs coding for the signal peptide and for the C-propeptide of type III procollagen were assembled as two fragments by using the full-length type III procollagen cDNA clone (H. Kuivaniemi, G. Tromp, unpublished results) as a template on specific primers in PCR (Saiki et al. 1985). To amplify sequences coding for the signal peptide, the upstream primer was KZ-35-III 5′-AAG ATC TAA AGA GTC TCA TGT CTG ATA TTT A-3′ that initiated at base number 76 from the start of transcription and that created a unique *Bgl*II site on the 5′ termini. The downstream primer was KZ-36-III 5′-AAT CGA TCT GTT GTG CCA AAA TAA TAG T-3′ that terminated at base number 177 and that created a unique ClaI site on its 5′ terminus. The primers were based on the published cDNA sequence (Ala-Kokko et al., 1989). The amplification conditions were: an initial denaturation at 94.5 °C for 2 min, followed by 30 cycles of denaturation at 94 °C for 15 sec, annealing at 60 °C for 15 sec and elongation at 72 °C for 15 sec, followed by a final elongation at 72 °C for 7 min. The PCR products were analyzed on 0.8% agarose gel. cDNA sequences coding for the complete C-propeptide and a portion of the 3′-untranslated region were amplified using the primer KZ-37-III 5′-ATT CGA TGA ACC AAT GGA TTT CAA AAT CAA C-3′ that initiated at position number 3766 from the start of transcription and created a unique *Cla*I site on the 5′-termini, and the primer KZ-III-18 5′-GGC GAA GCT TGT TGA CCT TTA CAG GAT GAA-3′ that terminated at base number 5337 in the 3′-untranslated region of the cDNA. The PCR conditions used were an initial denaturation at 94.5 °C for 30 sec, followed by 30 cycles of denaturation at 94 °C, annealing at 55 °C for 30 sec, and elongation at 72 °C for 90 sec followed by a final elongation at 72 °C for 7 min. The PCR product of 1586 bp contained a *Cla*I site engineered on KZ-37-III and a natural *Xba*I site internal to the PCR primer III-18 that were used to generate the intermediate construct coding for the type III signal-COOH propeptide. The PCR product was analyzed on a 1% agarose gel and cloned into the vector pCRII (Invitrogen). Individual clones were sequenced to ensure that no mutations were incorporated during the PCR. One clone containing the cDNA for the signal peptide and another clone containing the C-propeptide cDNA fragment were each digested with the restriction endonuclease *Cla*I (New England Biolabs) and ligated together into the *Bgl*II and *Xba*I in the vector

pCRII (Invitrogen) using T4 DNA ligase (USB) at 16 °C overnight (Sambrook et al. 1989). The resultant ligation products were amplified using the primers KZ-35-III and KZ-III-18 and PCR conditions identical to those used for the amplification of the C-propeptide (see above). The PCR product was analyzed on a 1% agarose gel, cloned into the vector pCRII (Invitrogen), and the sequence verified by DNA sequencing.

Construction of the transfer vector pVL-TyIII-3.1

The recombinant clone that contained both the signal peptide and C-propeptide was excised with the enzymes *BglI* (New England Biolabs) and *XbaI* (New England Biolabs), and subcloned into the same sites in the baculovirus transfer vector pVL1392 (Invitrogen). Restriction enzyme digestion and DNA sequencing were used to verify the orientation in the shuttle vector of the inserts of positive clones with respect to the polyhedron (polh) promoter (O'Reilly et al. 1992).

Co-transfection of transfer vector and viral DNA

Spodoptera frugiperda (Sf9) cells (Brown and Faulkner 1977) (3×10^6 cells per 25 cm^2 tissue culture flask) were co-transfected with pVL-TyIII-3.1 and *Autographica californica* nuclear polyhedrosis virus (AcNPV) DNA (Baculogold™; Pharmingen) using transfection buffer as specified by the supplier. The transfected cells were incubated in Grace's (non-supplemented) insect cell culture medium plus 10% fetal bovine serum at 27 °C for 6 days before the cell culture supernatant containing the extracellular recombinant baculovirus (rBac-TyIII-3.1) was collected and used to infect fresh Sf9 cells for virus amplification. High titer virus stocks of rBac-TyIII-3.1 (4×10^8 virus particles/ml) were obtained after two rounds of amplification (P2 and P3 virus) in which the recombinant baculovirus-infected cells was incubated at 27 °C for 4 days each in Sf900II serum-free insect cell culture medium (SFM; Gibco BRL Life Technologies, Inc.).

Expression of TyIII-3.1 recombinant protein in insect cells

Trichoplusia ni (5B1-4, High Five) cells (1×10^6 per ml) (Invitrogen Corp.) were infected with rBac-TyIII-3.1 at a multiplicity of infection of approximately 5 virus particles/cell and incubated in Sf900II SFM. The cell culture supernatants containing the TyIII-3.1 recombinant protein were analyzed by electrophoresis on 4 to 15% SDS-polyacrylamide gels (BioRad) and the protein bands stained with Coomassie blue (R250; Kodak) using standard methods (Sambrook et al. 1989).

Scale-up production of TyIII-3.1

Large quantities of TyIII-3.1 recombinant protein were produced in High Five Cells by infecting 1 L of suspension culture (1×10^6 cells per ml) with rBac-TyIII-3.1 virus in a 3 liter spinner flask fitted with a double paddle (Bellco Glass, Inc.). When the cell viability reached 90% (48 h p.i.), the cell culture supernatant was harvested by centrifugation in 250 ml bottles at 5000 × g for 10 min, and transferred to fresh bottles.

Purification of C- propeptide from media

p-Aminobenzamidine dihydrochloride acid (PABA; Sigma) to a final concentration of 1 mM, and N-ethylmaleimide (NEM; Sigma) to a final concentration of 10 mM were added to the cell culture supernatant containing the recombinant type III procollagen C- propeptide. The sample was filtered to remove cell debris and passed over a Con A Sepharose 4B matrix (Pharmacia) as described by Choglay et al. (1993). Bound proteins were eluted with 1 M methyl-α-D-manno-pyranoside and fractions collected. Samples from the fractions were separated by SDS-PAGE under both reducing and non-reducing conditions and products visualized by staining with Coomassie Brilliant Blue. All fractions that contained protein were pooled and dialyzed twice against 10 volumes of cation buffer (Pharmacia) containing 50 mM sodium acetate (pH 4.5), 5 mM NEM and 0.5 mM PABA. The proteins were loaded onto 25 ml of cation buffer equilibrated SP Sepharose Fast Flow cation matrix (Pharmacia) at a flow rate of 50–75 ml/h. The column was washed with 5–10 column volumes of cation buffer and the protein eluted with an NaCl gradient, 0 to 1 M in one liter. Five milliliter fractions were collected and absorbance at 280 μm measured. The C-propeptide eluted as a single peak. A sample of 100–500 ng of the protein was separated by electrophoresis on a 4 to 15% SDS-polyacrylamide gel (BioRad) and visualized by staining with Coomassie blue.

C- propeptide of type I procollagen was purified from chicken tendon cultures as described by Choglay et al. (1993).

Generation of cyanogen bromide fragments

Purified C-propeptides in 70% formic acid were reacted with cyanogen bromide using one small crystal per

100 µl of protein solution (about 1 mg) at room temperature for 14–18 h. The sample was lyophilized and resuspended in storage buffer (0.1 M Tris-HCl, 0.4 M NaCl, and 25 mM EDTA, pH 7.4) and separated by electrophoresis through a 4–15% gradient polyacrylamide gel in SDS under either reducing or non-reducing conditions. The proteins were visualized by either Coomassie blue or silver staining of the gel.

CD spectroscopic analysis

Circular dichroism (CD) spectra were recorded on an AVTV60DS spectrometer (AVIV Associates, Lakewood, NJ) equipped with a thermostated cell holder connected to a programmable circulating water bath (Lauda Model RC6 from Brinkmann Instruments, Westbury, NY). The instrument was calibrated by the two-point method (Chen and Yang, 1977). Sample concentrations were 7.4×10^{-5} g/ml for the chicken type I C-terminal propeptide and 5.7×10^{-5} g/ml for the human type III C-terminal propeptide as determined using the Bradford (1976) assay (BioRad, Hercules, CA). The samples were dialyzed into 10 mM Tris-HCl buffer and placed in 0.1 cm pathlength stoppered cells made of suprasil quartz (Hellma, Jamaica, NY). After the sample was equilibrated in the instrument at 20 °C for 30 min, the far UV CD spectrum from 250 nm to 195 nm was recorded. To assay thermal stability, the temperature was increased at a rate of 0.3 °C/min and the ellipticity at 220 nm was recorded every 5 °C between 20 °C and 60 °C. At 60 °C a final spectrum was recorded.

Results

Time course of protein induction

The recombinant virus, rBac-TyIII-3.3 contained the signal peptide of type III preprocollagen fused in-frame to the C-propeptide of type III procollagen. The fused fragments contained a new *Cla*I site and an additional codon for Ile at the C-terminus of the signal sequence. Baculovirus cells were infected with recombinant virus rBac-TyIII-3.1 and the media was removed from the cell culture after 24, 42, 48, 65 and 72 h post-infection. The proteins were separated on SDS-polyacrylamide gel electrophoresis under both reduced and nonreduced conditions. A highly induced band was initially seen in the media after 24 h and 42 h infection (Fig. 1). The total amount in media decreased as the cells began to lyse after 65 and 72 h. Therefore, there was no evidence of marked intracellular accumulation of the protein. As expected for the C-propeptide of type III procollagen, the recombinant protein was a trimer under nonreduced conditions with a molecular weight of approximately 100 kDa (not shown). After reduction, it was a monomer of approxi-

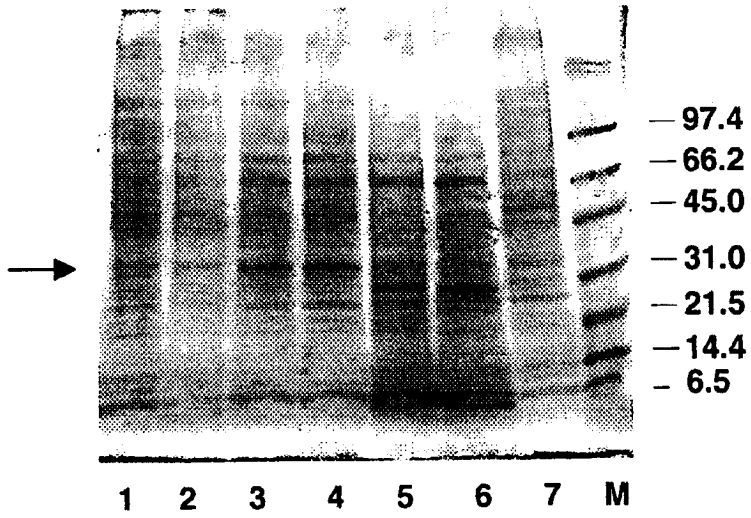

Figure 1. Time course of recombinant C-propeptide of proα1(III) collagen expression using a baculovirus system. Samples of medium were reduced and electrophoresed through a 4–15% SDS-polyacrylamide gel and proteins stained using Coomassie Blue. Lanes 1 and 7, total protein from uninfected cells. Lanes 2–6, total protein 24, 42, 48, 65, and 72 h post-infection. Arrow indicates expressed C-propeptide. M, molecular weight markers (BioRad Broad Range) with values in kDa.

mately 32 kD. After incubation of the cells for 65 and 72 h, the recombinant monomer had a molecular weight of about 30 kDa under reducing conditions, an observation suggesting degradation of the recombinant protein at the later time points (Fig. 1). Addition of protease inhibitors during cell culture had little effect in preventing degradation of the recombinant protein.

Purification of recombinant material from insect cell culture

Recombinant C-propeptide of type III procollagen secreted into the culture medium was purified by chromatography on a concanavalin A column and then on a cation exchange column. The bound protein was eluted from the cation exchange column using a salt gradient and the C-propeptide of proα1(III) collagen eluted at 0.15 M NaCl. The fractions were analyzed by SDS polyacrylamide gel electrophoresis (Fig. 1). The typical yields of purified protein were 1 to 2.5 mg per liter of media.

Characterization of inter- and intra-chain disulfide bonds by cyanogen bromide cleavage

The recombinant protein was assayed for presence of disulfide bonds by cyanogen bromide mapping (Fig. 3). The cysteines numbered C_1 to C_4 that lie closest to the triple helix of the molecule, have been shown to be involved in inter-chain disulfide linkages (Koivu, 1987; Lees and Bulleid, 1994). In unreduced conditions they should generate a disulfide-linked fragment of 20.5 kDa. The C-terminal cysteines C_5 to C_8 form intra-chain disulfide linkages. Under non-reducing conditions they should generate a fragment of 17.4 kDa. The recombinant protein was cleaved with cyanogen bromide and separated by SDS-PAGE under nonreducing or reducing conditions. As expected, the largest peptides obtained under non-reducing conditions were about 20 kDa and 17 kDa (arrows A and B in Fig. 3B). The molecular weights of the reduced peptides were more difficult to relate to the expected sizes apparently because of incomplete cleavage by cyanogen bromide.

Analysis of protein structure by circular dichroism spectroscopy

The spectrum showed a broad peak of negative ellipticity in the wave-length range of 200 to 240 nm for contributions of known secondary structure elements (Fig. 4A). Analysis of the spectrum by the algorithm of Chen et al. (1972) indicated that the trimer consists of

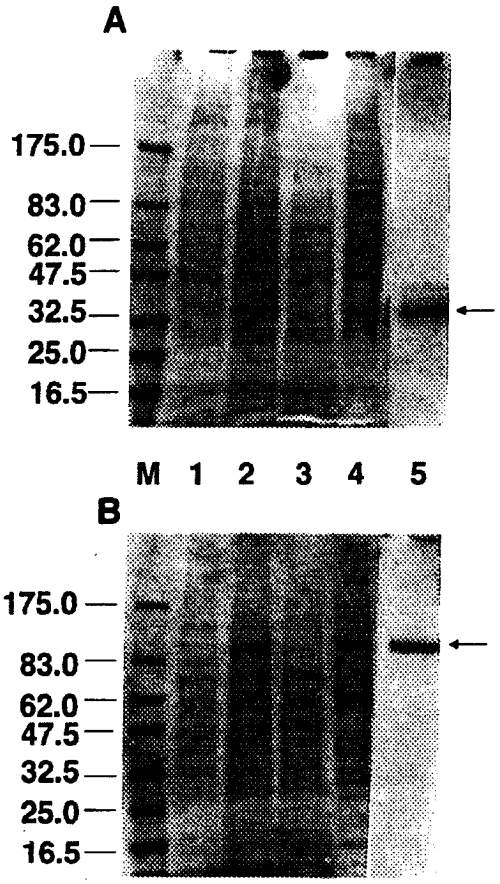

Figure 2. Purification of recombinant C-propeptide of proα1(III) collagen from insect cell cultures. Samples were electrophoresed through a 4–15% SDS-polyacrylamide gel and proteins stained using Coomassie Blue. Panel A, reducing conditions; Panel B, non-reducing conditions. Lane 1, media from uninfected cells as negative control; Lane 2, total media 48 h post-infections; Lane 3, void volume from concanavalin column; Lane 4, bound proteins from concanavalin column eluted with methyl-a-D-mannopyranoside; Lane 5, bound protein eluted with salt gradient from SP Sepharose cation column. M, molecular weight marker (New England Biolabs) with values in kDa. Arrow indicates expressed C-propeptide.

Table I. CD spectra of human recombinant type III and chick type I C-propeptides.

	α-helix	α-sheet	α-turn	Random
Human type III	7%	31.7%	27.8%	33.3%
Chick type I	7.7%	39.8%	20.8%	31.4%

Values from Spectra (Fig. 4A) were normalized to 100%.

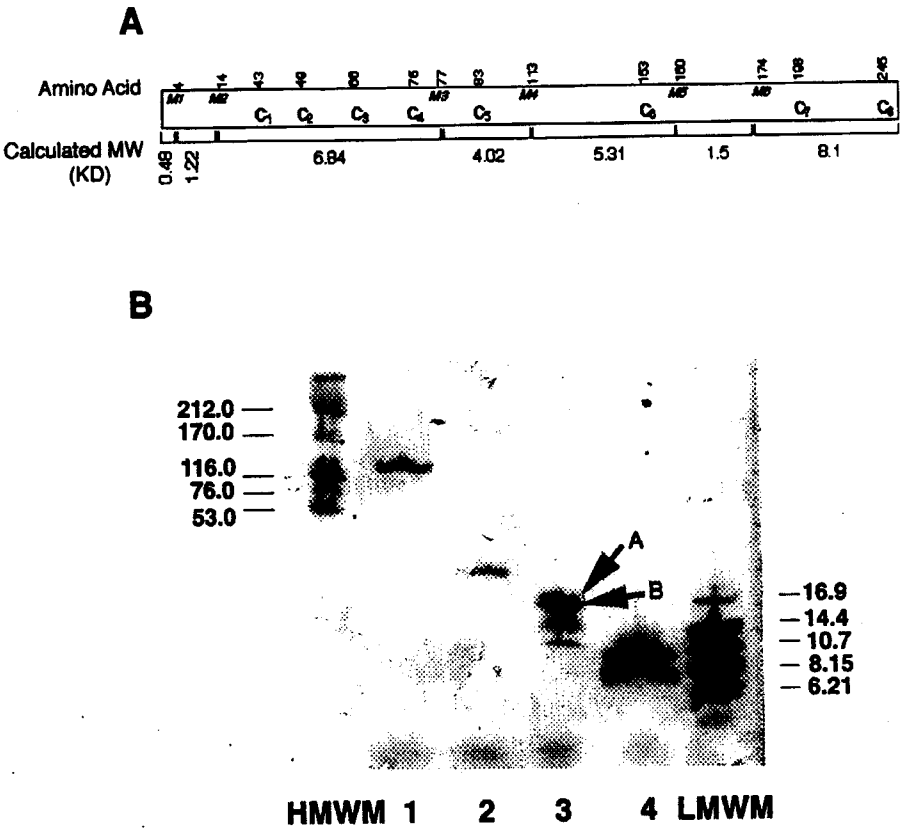

Figure 3. A. Diagrammatic illustration of expected sizes of fragments after cyanogen bromide cleavage. Top panel represents a single chain of the C-propeptide from proα(III) collagen. Numbers at top represent amino acid positions; numbers at the bottom represent calculated molecular weights of each fragment after cleavage with cyanogen bromide; M_{1-6} represents methionine residues; C_{1-8} represent cysteine residues.
B. Inter- and intra-chain disulfide bond characterization by cyanogen bromide cleavage. Lane HMWM: High molecular weight markers (Pharmacia). Lane 1: Unreduced recombinant C-propeptide. Lane 2: Reduced recombinant C-propeptide. Lane 3: Cyanogen bromide fragments of C-propeptide, unreduced. Lane 4: Cyanogen bromide digest of C-propeptide, reduced. Lane LMWM: Low molecular weight markers (Pharmacia) in kDa. The arrows A and B indicate fragments with intrachain and interchain disulfide bonds, respectively.

7.0% α-helix, 31.7% β-sheet, 27.8% β-turn and 33.3% random structure (Table 1). Since the amino acid composition of the human type III C-propeptide (Ala-Kokko et al 1989) is highly homologous to the chicken type I collagen C-propeptide (Bernard et al. 1983a,b), the secondary structure of the C-propeptide of chicken type I procollagen obtained from organ cultures of tendons was simultaneously determined. The data indicated that the trimer of type I procollagen C-propeptides had about the same structure as the recombinant type III C-propeptide (Table 1).

The CD spectra obtained at 20 °C and at 60 °C for the two C- propeptides were qualitatively similar. There was a slight decrease in the relative amount of α-helix at 60 °C accompanied by a corresponding increase in unordered structures (not shown). The results, however, were well within the ±5% limits of the method. As a more direct measure of conformational stability, the data in Figure 4B were obtained by monitoring the ellipticity at 220 nm on a single sample of each propeptide as the temperature was slowly increased from 20 °C to 60 °C. This plot indicated that the human recombinant type III C-propeptide was more stable than the chick type I C-propeptide. Comparisons of the full spectra at 60 °C (not shown) with the 20 °C spectra supported the same conclusion.

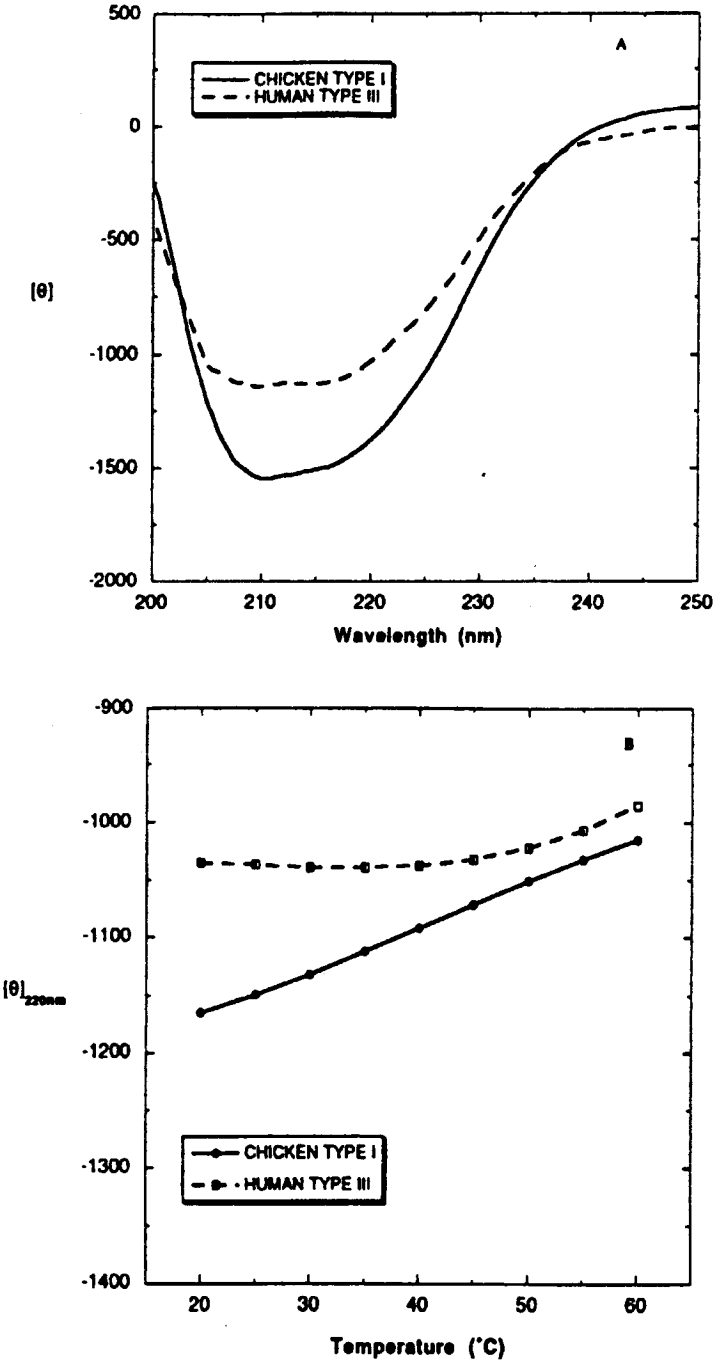

Figure 4. CD spectroscopic analysis of C-propeptide of chicken type I and human recombinant type III procollagens. A, CD spectrum of the propeptides at 20 °C. The data are expressed in terms of the mean residue ellipticity, [θ]. B, A plot of [θ]220nm as a function of temperature between 20 °C and 60 °C.

Discussion

The results here demonstrate that the C-propeptide of human type III procollagen can be synthesized in a baculovirus recombinant system. The protein was secreted into the medium, and the yield (1 to 2.5 mg per liter) was adequate for a number of biological studies. Peptide mapping (Fig. 3) indicated that the recombinant protein contained the correct disulfide bonds as defined by Koivu (1987) and Lees and Bulleid (1994). Also, the protein had a conformation similar to that of the native C-propeptide of type I procollagen as assayed by CD. However, the type III C-propeptide was slightly more stable to heat denaturation, perhaps because it is a homotrimer and the type I C-propeptide is a heterotrimer.

Ten mutations altering the structure of the C-propeptide of type I procollagen were found in patients with osteogenesis imperfecta (see Kuivaniemi et al., 1997). The phenotypes range from mild to lethal, but the consequences of most of the mutations for the functions of the protein have not been defined. The development of recombinant technologies to study the structure and the thermal stability of the C-propeptides will facilitate the correlation of the phenotype and mutations in specific domains within the protein.

Of special interest is that two previous reports (Tomita et al., 1995; Lamberg et al., 1996) described synthesis of full-length type III procollagen in the baculovirus system. The recombinant protein formed the correct disulfide bonds, but it was retained intracellularly and no significant amounts were secreted. Therefore, it was difficult to purify the protein without resorting to protease digestions that degrade the N- and C-propeptides. Recovery of the type III C-propeptide from the medium under the conditions described here is probably explained by the smaller size of the protein and perhaps a lesser tendency to aggregate.

References

Ala-Kokko, L, Kontusaari, S., Baldwin, C., Kuivaniemi, H. and Prockop, D.J.: Structure of cDNA clones coding for the entire preproα1(III) chain of human type III procollagen. Differences in protein structure from type I procollagen and conservation of codon preferences. *Biochem. J. 260*: 508–517, 1989.

Aycock, R.S., Raghow, R., Stricklin, G.P., Seyer, J.M. and Kang, A.H.: Post-transcriptional inhibition of collagen and fibronectin synthesis by a synthetic homolog of a portion of the carboxy-terminal propeptide of human type I collagen. *J. Biol. Chem. 261*: 14355–14360, 1986.

B_chinger, H.P., Bruckner, P., Timpl, R., Prockop, D.J. and Engel, J.: Folding Mechanism of the triple helix in type III collagen and type III pN-collagen. *Eur. J. Biochem. 106*: 619–632 (1980)

Bernard, M.P., Chu, M.L., Myers, J.C., Ramirez, F., Eikenberry, E.F. and Prockop, D.J.: Nucleotide sequences of complementary deoxyribonucleic acids for the proα1 chain of human type I procollagen. Statistical evaluation of structures that are conserved during evolution. *Biochemistry 22*: 5213–5223, 1983a.

Bernard, M.P., Myers, J.C., Chu, M.L., Ramirez, F., Eikenberry, E.F. and Prockop, D.J.: Structure of a cDNA for the proα2 chain of human type I procollagen. Comparison with check cDNA for proα2(I) identifies structurally conserved features of the protein and the gene. *Biochemistry 22*: 1139–1145, 1983b.

Bradford, M.M.: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*: 248–254, 1976.

Brown, M. and Falkner, P.: A plaque assay for nuclear polyhedrosis viruses using a solid overlay. *J. General Virology 36*: 361–364, 1977.

Chen, G.C. and Yang, J.T.: Two point calibration of circular dichrometer with d-10-camphorsulfonic acid. *Anal. Lett. 10*: 1195–1207, 1977.

Chen, Y.-H., Yang, J.T. and Martinez, H.M.: Determination of the secondary structures of proteins by circular dichroism and optical rotary dispersion. *Biochemistry 11*: 4120–4131, 1972.

Chipman, S.D., Sweet, H.O., McBride, J.D.J., Davisson, M.T., Marks, J.S.C., Shuldiner, A.R., Wenstrup, R.J., Rowe, D.W. and Shapiro, J.R.: Defective proα2(I) collagen synthesis in a recessive mutation in mice. A model of human osteogenesis imperfecta. *Proc. Natl. Acad. Sci. 90*: 1701–1705, 1993.

Choglay, A.A., Purdom, I.F. and Hulmes, D.J.: Procollagen binding to sphingomyelin. *J. Biol. Chem. 268*: 6107–6114, 1993.

Deak, S.B., Nicholls, A., Pope, F.M. and Prockop, D.J.: The molecular defect in a nonlethal variant of osteogenesis imperfecta. Synthesis of proα2(I) chains which are not incorporated into trimers of type I procollagen. *J. Biol. Chem. 258*: 15192–15197, 1983.

Dion, A.S. and Myers, J.C.: COOH-terminal propeptides of the major human procollagens. Structural, functional and genetic comparisons. *J. Molec. Biol. 193*: 127–143, 1987.

Halila, R. and Peltonen, L.: Purification of human procollagen type III N-proteinase from placenta and preparation of antiserum. *Biochem. J. 239*: 47–52, 1986.

Hojima, Y., McKenzie, J., van der Rest, M. and Prockop, D.J.: Type I procollagen N-proteinase from chick embryo tendons. Purification of a new 500-kDa form of the enzyme and identification of the catalytically active polypeptides. *J. Biol. Chem. 264*: 11336–11345, 1989.

Hojima, Y., van der Rest, M. and Prockop, D.J.: Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization. *J. Biol. Chem. 260*: 15996–16003, 1985.

Katayama, K., Seyer, J.M., Raghow, R. and Kang, A.H.: Regulation of extracellular matrix production by a chemically synthesized subfragment of type I collagen carboxy peptide. *Biochemistry 30*: 7097–7104, 1991.

Kessler, E., Takahara, K., Biniaminov, L., Brusel, M. and Greenspan, D.S.: Bone morphogenic protein-1: The type I procollagen C-proteinase. *Science 271*: 360–362, 1996.

Koivu, J.: Identification of disulfide bonds in carboxy-terminal propeptides of human type I. *FEBS Letts. 212*: 229–232, 1987.

Kuivaniemi, H., Tromp, G. and Prockop, D.J.: Mutations in fibrillar collagens (types I, II, III and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. *Hum. Mut. 9*: 300–315, 1997.

Lamande, S.R., Cressler, S.D., Golub, S.B., Byers, P.H., Chan, D., Cole, W.G., Sillence, D.O. and Bateman, J.F.: Endoplasmic reticulum-mediated quality control of type I collagen production by cells from osteogenesis imperfecta patients with mutations in the proα1(I) chain carboxyl-terminal propeptide which impair subunit assembly. *J. Biol. Chem. 270*: 8642–8649, 1995.

Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T. and Kivirikko, K.I.: Characterization of human type III collagen expressed in a baculovirus system. Production of a protein with a stable triple helix required coexpression with the two types of recombinant prolyl 4-hydroxylase subunit. *J. Biol. Chem. 271*: 11988–11995, 1996.

Lees, J.F. and Bulleid, N.J.: The role of cysteine residues in the folding and association of the COOH-terminal propeptide of type I and III procollagen. *J. Biol. Chem. 269*: 24354–24360, 1994.

Li, S.-W., Sieron, A.L., Fertala, A., Hojima, Y., Arnold, W.V. and Prockop, D.J.: The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. *Proc. Natl. Acad. Sci. USA 93*: 5127–5230, 1996.

Nicholls, A.C., Pope, F.M. and Schloon, H.: Biochemical heterogeneity of osteogenesis imperfecta. New variant. *Lancet 1*: 1193, 1979.

O'Reilly, D.R., Millder, L.R. and Lucknow, V.A.: Baculovirus expression vectors: A laboratory manual. Vol. 1. New York: W. H. Freeman & Co., 1992.

Prockop, D.J. and Hulmes, D.J.S.: Assembly of collagen fibrils *de novo* from soluble precursors: Polymerization and copolymerization of procollagen, pN-collagen, and mutated collagens. In *Extracellular Matrix Assembly and Structure*, Yurchenco, P.D., Birk, D.E., Mechan, R.P., eds., Academic press, New York, pp. 47–90, 1994.

Prockop, D.J. and Kivirikko, K.I.: Collagens: Molecular biology, diseases and potentials for therapy. *Ann. Rev. Biochem. 64*: 403–434, 1995.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N.: Enzymatic amplification of b-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230*: 1350–1354, 1985.

Sambrook, J., Fritsch, E.F., Maniatis, T.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989.

Schofield, J.D., Uitto, J. and Prockop, D.J.: Formation of interchain disulfide bonds and helical structure during biosynthesis of procollagen by embryonic tendon cells. *Biochemistry 13*: 1801–1806, 1974.

Tomita, M., Ohkura, N., Ito, M., Kato, T., Royce, P.M. and Kitajima, T. Biosynthesis of recombinant human proα1(III) chains in a baculovirus expression system: production of disulphide-bonded and non-disulphide-bonded species containing full-length triple helices. *Biochem. J. 312*: 847–853, 1995.

Uitto, J. and Prockop, D.J.: Incorporation of proline analogues into collagen polypeptides. Effects on the production of extracellular procollagen and on the stability of the triple-helical structure of the molecule. *Biochim. Biophys. Acta 336*: 234–251, 1974a.

Uitto, J. and Prockop, D.J.: Intracellular hydroxylation of nonhelical protocollagen to form triple-helical procollagen and subsequent secretion of the molecule. *Eur. J. Biochem. 43*: 221–230, 1974b.

Weston, S.A., Hulmes, D.J.S., Mould, A.P., Watson, R.B. and Humphries, M.J.: Identification of integrin a21b1 as cell surface receptor for the carboxy-terminal propeptide of type I procollagen. *J. Biol. Chem. 269*: 20982–20988, 1994.

Wu, C.H., Donovan, C.B., and Wu, G.Y.: Evidence for pretranslational regulation of collagen synthesis by procollagen propeptides. *J. Biol. Chem. 261*: 10482–10484, 1986.

Wu, C.H., Walton, C.M. and Wu, G.Y.: Propeptide-mediated regulation of procollagen synthesis in IMR-90 human lung fibroblast cell cultures. Evidence for transcriptional control. *J. Biol. Chem. 266*: 2983–2987, 1991.

Darwin J. Prockop, Center for Gene Therapy, Allegheny University of the Health Sciences, Philadelphia, PA 19102

Received June 5, 1997; accepted July 17, 1997