Published in final edited form as:

Protein Expr Purif. 2007 April ; 52(2): 403–409. doi:10.1016/j.pep.2006.10.016.

Expression, Purification, and Refolding of Recombinant Collagen $\alpha 1(XI)$ Amino Terminal Domain Splice Variants

Lisa R. Warner¹, Christina M. Blasick², Raquel J. Brown², and Julia Thom Oxford^{1,2,*}

¹Department of Materials Science and Engineering, Boise State University, Boise, Idaho

²Department of Biology and Biomolecular Research Center, Boise State University, Boise, Idaho

Abstract

The amino terminal domain of collagen type XI $\alpha 1$ chain is a noncollagenous structure that is essential for the regulation of fibrillogenesis in developing cartilage. The amino terminal domain is alternatively spliced at the mRNA level, resulting in proteins expressed as splice variants. These splice variants, or isoforms, have unique distribution in growing tissues, alluding to distinct roles in development. We report here a rapid and straightforward method for expression, purification and *in vitro* folding of recombinant collagen XI isoforms $\alpha 1(XI)$ NTD[p7] and $\alpha 1(XI)$ NTD[p6b+7]. The recombinant isoforms were expressed in *Escherichia coli* as bacterial inclusion bodies. Unfolded carboxy terminal polyhistidine tagged proteins were purified via nickel affinity chromatography and refolded with specific protocols optimized for each isoform. Purity was assessed by SDS-PAGE and correct secondary structure by a comparison of circular dichroism data with that obtained for Npp. Protein expression and purification of the recombinant collagen XI splice variants will allow further studies to elucidate the structure and molecular interactions with components of the extracellular matrix. This research will clarify the mechanism of collagen XI mediated regulation of collagen fibrillogenesis.

Introduction

The collagens of articular cartilage provide tensile strength by forming an elaborate scaffold of covalently cross-linked heterotypic collagen fibrils. Collagen fibrils of cartilage are composed of approximately 80% type II, 10% type IX, and 10% type XI [1]. Fibrillar collagen type XI is detected on the surface of these heterotypic collagen fibrils and plays a crucial role in regulating type II collagen assembly [2] and [3]. Blaschke et. al. determined that the diameter of cartilage collagen fibrils is directly dependent on the ratio of collagen type XI to collagen type II [4]. Collagen XI also plays a critical regulatory role in extracellular matrix organization in early development as demonstrated by the chondrodystrophic mouse (*cho/cho*) in which collagen fibrils are large in diameter and the growth plates are disorganized [3].

The type XI collagen molecule is a trimer consisting of the polypeptides $\alpha 1(XI)$, $\alpha 2(XI)$, and $\alpha 3(XI)$; where $\alpha 3(XI)$ is found as an over-hydroxylated version of $\alpha 1(XI)$. All three α chains include non-triple helical domains at both the carboxyl and amino termini and an interrupted

*To whom correspondence should be addressed, Corresponding author: Julia Thom Oxford, Ph.D, Department of Biology, Biomolecular Research Center, 1910 University Drive, MS 1515, Boise State University, Boise, Idaho 83725, Telephone: 208-426-2395, Fax: 208-426-4267, Email: E-mail: joxford@boisestate.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

minor helix [5] and [6]. Although similar in structure to other amino propeptides of the fibrillar collagen chains $\alpha 2(\text{XI})$, $\alpha 1(\text{V})$, and $\alpha 3(\text{V})$, the amino terminal domain (NTD) of $\alpha 1(\text{XI})$ is retained on the surface of collagen fibrils with a relatively long half-life [7]. It has been postulated that regulation of type II collagen is managed via the NTD of $\alpha 1(\text{XI})$ by sterically hindering assembly of collagen type II, yet the exact mechanistic details are unclear [4].

The $\alpha 1(\text{XI})$ amino propeptide (Npp) shown in Fig. 1, common to all isoforms, is encoded by exons 1 through 5 and contains a bone morphogenic protein 1 (BMP-1) processing site at AAQA/QHIDE previously identified [8] and [9]. The BMP-1 processing site is located seven amino acids from the carboxyl end of the polypeptide region encoded by exon 5 and defines the carboxyl terminus of the $\alpha 1(\text{XI})$ Npp. The variable region, which spans exons 6a, 6b, 7, 8 and 9, encodes several biochemically unique domains (Fig. 1A) which are expressed in a manner that includes alternative splicing at the mRNA level, resulting in expression of different protein splice variants [10]. The most abundant forms of the $\alpha 1$ chain of type XI collagen are $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] comprising 40% and 30% of expressed $\alpha 1(\text{XI})$ respectively (Fig. 1B) [11]. Isoforms of the pro- $\alpha 1(\text{XI})$ gene exhibit tissue-dependent temporal and spatial expression in rat fetal cartilage [12] and [13]. The $\alpha 1(\text{XI})$ NTD[p7] is expressed ubiquitously in developing cartilage while $\alpha 1(\text{XI})$ NTD[p6b+7] is specifically expressed at the leading edge of the growth plate and precedes the formation of the bony collar [12].

Due to the prevalence of the $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] isoforms, we have chosen to focus our research on these proteins. In order to further characterize the structure and function of $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7], it is necessary to produce large quantities of these isoforms in pure, soluble and properly folded form. Here we describe a straightforward and efficient process for the production of recombinant collagen XI $\alpha 1(\text{XI})$ NTDs from *Escherichia coli* cytosolic inclusion bodies. We report the construction of synthetic pro- $\alpha 1(\text{XI})$ genetic constructs encoding splice variants, expression of recombinant collagen XI $\alpha 1(\text{XI})$ NTD isoforms in *Escherichia coli*, purification via histidine-linked nickel affinity chromatography under denaturing conditions, *in vitro* refolding to generate pure, soluble proteins and analysis of resulting structural characteristics.

Materials and Methods

Materials

Unless otherwise noted, reagents were purchased from Fisher Scientific (Hampton, NH).

Cloning and Expression of $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7]

The cDNA encoding the $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] isoforms of collagen $\alpha 1(\text{XI})$ chain from rat were independently cloned into expression vector pET11a (Stratagene, La Jolla, CA) as previously described [14] using PCR products obtained with an upstream primer which included a unique NheI restriction site [14] and a downstream primer encoded six histidine residues, a stop codon, a SpeI site, and a unique BamHI site at the end [14]. After amplification, the sequences were cloned into the pET 11a expression vector, and the identity verified by sequencing. After transformation into *E. Coli* BL21(DE3) competent cells, cultures were grown overnight at 37 °C with shaking at 250 rpm and 1 ml was used to inoculate 100 ml of Luria-Bertani (LB) broth (Amersham Biosciences, Piscataway, NJ) supplemented with 50 $\mu\text{g/ml}$ carbenicillin. Protein expression was induced with 0.4 mM isopropyl- β -D-thiogalactopyranoside (IPTG, Roche Applied Science, Indianapolis, IN) at an A_{600} of 0.4 - 0.6. Cells were harvested 5 h post induction by centrifugation at 15,000g for 30 min.

Extraction and Purification of $\alpha 1(XI)$ NTD [p7] and $\alpha 1(XI)$ NTD[p6b+7]

Cells were lysed and proteins solubilized using bacterial protein extraction reagent (BPER, Pierce Biotechnology, Rockford, IL) with 0.1 unit/ml of DNase (Roche Applied Science, Indianapolis, IN) and 3 μ g/ml lysozyme. The insoluble fraction was washed five times in 50 mM Tris-HCl, pH 7.65 and the remaining pellet was collected by centrifugation at 15,000g for 30 min at 4° C. The isolated inclusion bodies were resuspended in a solubilizing buffer (SB) (70 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$, pH 7.65, containing 6 M guanidine hydrochloride (GuHCl), 14 mM β -mercaptoethanol, and 80 mM NaCl) and incubated at room temperature for 1 h. The solution was centrifuged at 15,000g for 30 min at 4° C and insoluble debris was removed. The supernatant was incubated with Ni-NTA Superflow (Qiagen) overnight at 4° C with stirring. Purification was performed using nickel affinity chromatography on a BioLogic LP chromatography system (Bio-Rad, Hercules, CA). Non-specifically bound proteins were removed by exhaustive washing with SB at a flow rate of 1 ml/min. Bound proteins were eluted with 200 mM imidazole in SB and collected into 1 ml fractions. Absorbance at 280 nm was measured and recorded using BioLogic LP Data View Software, version 1.01 (Bio-Rad, Hercules, CA) throughout purification. The fractions were evaluated under reducing conditions by SDS-PAGE. Gels were stained with Bio-Safe Coomassie brilliant blue G250 dye (Bio-Rad, Hercules, CA) and qualitatively evaluated for purity.

Denaturant Removal/Protein Refolding

Refolding was performed using conditions optimized for each of the recombinant proteins. Refolding of $\alpha 1(XI)$ NTD[p7] was initiated by diluting purified protein into 20 volumes of 55 mM Tris-HCl, pH 8.2 containing 21 mM NaCl, 0.88 mM KCl, 2 mM reduced glutathione, 4 mM oxidized glutathione, and 0.1 mM EDTA. Refolding of $\alpha 1(XI)$ NTD[p6b+7] was initiated under the same conditions but with the addition of 0.88 M L-arginine to the buffer. Protein was added in five equal increments at two minute intervals on ice. Refolding reactions were allowed to proceed overnight at 4° C. Refolded proteins were dialyzed exhaustively into 50 mM Tris-HCl, pH 7.5 containing 100 mM NaCl, for subsequent experiments.

Spectral Analysis and Quantification

Proteins were analyzed both by evaluating the absorbance at 280 nm and the Bradford assay (Bio-Rad, Hercules, CA). Baselines were corrected for each sample to account for buffer content. Molar extinction coefficients were calculated using ProtParam [15].

Secondary Structure Analysis using Circular Dichroism

Circular dichroism (CD) spectra were monitored on a Jasco-810 instrument. Proteins were dialyzed exhaustively against 20 mM NaH_2PO_4 , pH 7.5 containing 100 mM NaF, and concentrated with a Millipore Amicon Ultra-15, NMWL 10,000 centrifugal concentrator. Samples were examined in a circular demountable 0.01 mm pathlength Supracil cell (Hellma, USA, Ltd., Plainview, NY). The instrument was calibrated using D (+) camphor sulfonic acid (CSA) at 192 nm and 290 nm. Three spectra and three baselines were collected at 0.5 nm intervals from 270 nm to 178 nm at 20° C. The high tension (HT) never exceeded 500 V.

CD spectra were processed using CDtool software [16]. The raw spectra were averaged, contributions from buffer were subtracted, and the resulting spectra smoothed using a Savitsky-Golay filter [17] with a smoothing window of seven. The calculated change in extinction coefficient ($\Delta\epsilon$) values was determined using the mean residue weight value which was calculated from its amino acid sequence and calibrated to CSA.

Secondary structure analyses were performed using the DICHROWEB Web server [18] and [19] using the CONTINLL algorithm [20] and [21]. Reference dataset one was used in fitting

[22]. The normalized root-mean-square deviation (NRMSD) parameter was used to evaluate the goodness of fit of calculated data to experimental data [23]. It is defined as $\sum[(\theta_{\text{exp}} - \theta_{\text{calc}})^2/(\theta_{\text{exp}})^2]^{1/2}$ over all wavelengths, where θ_{exp} and θ_{calc} are the experimental ellipticities and the back-calculated ellipticities of the spectra for the predicted structure.

Results

Expression and Extraction of $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7]

The level of protein expression was evaluated by comparing pre- and post-induction samples of the pET11a-NTD[p7] and pET11a-NTD[p6b+7] by SDS-PAGE. Results indicated an increased level of expression of protein bands consistent with the molecular weights of $\alpha 1(\text{XI})$ NTD[p7] (45 kDa apparent molecular weight) and $\alpha 1(\text{XI})$ NTD[p6b+7] (50 kDa apparent molecular weight) as shown in Figs 2A and 2B respectively.

Extraction of soluble cellular *E. coli* proteins from insoluble inclusion bodies is shown in Figs. 2C and 2D. Insoluble *E. coli* proteins and recombinant isoforms were found in the suspensions while the soluble cellular *E. coli* proteins were washed away in the supernatants.

Purification of Recombinant Proteins

Recombinant isoforms, $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7], present in inclusion bodies were solubilized in 6 M GuHCl and purified on a nickel affinity column. The absorbance at 280 nm of the eluted solutions was recorded and representative resulting chromatograms for $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] are shown in Fig. 3. Fractions were evaluated for each purification under reducing conditions by SDS-PAGE and those containing $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] are shown in Fig. 3, lower panels. Fractions containing the recombinant protein were pooled for subsequent refolding and dialysis. Purification yielded 0.250 mg and 0.126 mg for $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] from 100 ml cultures respectively (Table 1). Apparent molecular weights for NTD[p7] and NTD[p6b+7] are 45 and 50 kDa respectively (calculated molecular weights are 30.5 kDa and 36.3 kDa for NTD[p7] and NTD[p6b+7] respectively).

Protein Refolding

The $\alpha 1(\text{XI})$ NTD[p7] and the $\alpha 1(\text{XI})$ NTD[p6b+7] were refolded as described in Materials and Methods. The $\alpha 1(\text{XI})$ NTD[p6b+7] was refolded with the addition of 0.88 M L-arginine to the buffer to help reduce aggregation [24]. SDS-PAGE gels of refolded, dialyzed and concentrated approximately ten fold $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] are shown in Figs. 4A and B respectively. Refolding yielded 0.192 mg and 0.115 mg for $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] from 100 ml cultures respectively (Table 1).

Secondary Structure Analysis using Circular Dichroism

Refolded proteins demonstrate secondary structures similar to the $\alpha 1(\text{XI})$ Npp, as shown in Fig. 5 and Table 2 [14]. Spectral data obtained for $\alpha 1(\text{XI})$ NTD[p7] demonstrated an increase in both α helical content as well as β sheet, while β turn and unordered content was lower than $\alpha 1(\text{XI})$ Npp alone [14]. Fig. 5 shows an overlay of the resulting spectra with results summarizing the secondary structural features in Table 2. Spectral differences between $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] are characteristic of the additional amino acids found in the variable region p6b, and include both α helix and β sheet.

Discussion

Type XI collagen is necessary for the successful assembly and structural integrity of fetal cartilage [26] and [3]. It is a quantitatively minor collagen which appears to regulate collagen

type II fibrillogenesis in building an extensive network of thin collagen fibrils [3]. *In vitro* studies suggest that the regulatory behavior of type XI collagen is dependent on its complex amino terminal domain (NTD) that may be proteolytically removed, albeit at a relatively slow rate [3], [4] and [7]. The $\alpha 1(\text{XI})$ NTD is a globular domain with similarities to the thrombospondin 1 and LNS (laminin, neurexin, and sex hormone binding) domains [27]. Retention of the intact or partially processed NTD at the surface of collagen II fibrils likely restricts lateral growth by steric hindrance and may contribute to structural integrity of cartilage by providing stabilizing interactions between collagen fibrils and/or other extracellular matrix components [2].

Isoforms of the $\alpha 1(\text{XI})$ chain have been shown, by immunohistological and immunofluorescence detection, to distribute in a temporal, spatial and tissue specific manner with $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] found primarily in cartilage [11] and [7]. During development of long bones and cartilage in rats, $\alpha 1(\text{XI})$ NTD[p7] is uniformly expressed while $\alpha 1(\text{XI})$ NTD[p6b+7] is restricted to the lateral boundaries of the diaphyseal cartilage, radius and ulna [12]. Just prior to ossification, $\alpha 1(\text{XI})$ NTD[p6b+7] is detected only in the diaphysis directly adjacent to the perichondrium, implicating a possible role in bone mineralization [12].

The task of elucidating the function of $\alpha 1(\text{XI})$ NTD splice variants relies on the consumption of large quantities of protein relative to the native concentration found in tissues. An overexpression system was chosen to produce greater quantities for biophysical characterization and functional assays. The majority of recombinant proteins were expressed in inclusion bodies which required refolding in addition to purification. Four cysteine residues are involved in the formation of two disulfide bonds and have been shown to form efficiently and specifically with stability under similar conditions of refolding as demonstrated by liquid chromatography-tandem mass spectrometry of tryptic peptides in the absence and presence of reducing agent [14].

Previous studies indicate that the $\alpha 1(\text{XI})$ NTD is primarily β -strand, β -turn and unordered [14] and [27]. Proteolytic removal of the NTD fragment from the collagen molecule occurs at the BMP-1 processing site with a $t_{1/2}$ of 18 h after synthesis [25] and [8]. In tissue, the bulk of soluble and therefore extractable NTD is found as the cleaved form of $\alpha 1(\text{XI})\text{Npp}$, while the $\alpha 1(\text{XI})$ NTD[p7] and $\alpha 1(\text{XI})$ NTD[p6b+7] remain on the surface of the collagen fibril, and are not easily extractable from tissue. This precludes isolation of intact isoforms from native tissue, so comparisons are made to $\alpha 1(\text{XI})$ Npp [14]. The $\alpha 1(\text{XI})$ NTD[p7] incorporates an additional 66 amino acids (aa) to the 223 found in $\alpha 1(\text{XI})$ NTD. The data from fitting refolded $\alpha 1(\text{XI})$ NTD[p7] CD spectra show an approximate 7% increase in α -helical content in the $\alpha 1(\text{XI})$ NTD[p7] when compared to the $\alpha 1(\text{XI})$ NTD. This may be due to the extended 66 aa "tail" for which the exact structure is yet to be determined. The $\alpha 1(\text{XI})$ NTD[p6b+7] has an additional 117 aa extension which CD spectra analysis suggest increases the ordered structure content, with respect to both α -helix and β -strand. This is in contrast to the $\alpha 1(\text{XI})$ NTD[p6a+8] splice variant reported previously, which demonstrated a decrease in content of ordered structure when compared to $\alpha 1(\text{XI})$ Npp [14]. Exon 6b is rich in codons for the positively charged amino acids arginine and lysine, which significantly alter the theoretical isoelectric point (pI) from 5.75 ($\alpha 1(\text{XI})$ NTD[p7]) to 9.19 ($\alpha 1(\text{XI})$ NTD[p6b+7]) [28], [10], and [13]. Proteins high in β -strand and β -turns are often difficult to denature and refold effectively without aggregation. The addition of L-arginine was useful in preventing aggregation during refolding of $\alpha 1(\text{XI})$ NTD[p6b+7] as a "neutral crowder" but some aggregation did occur during dialysis [24]. Previous studies showed a possible multimerization within the $\alpha 1(\text{XI})$ NTD and the additional charged tail afforded from the peptide encoded by the 6b exon could cause heightened self association, leading to aggregation if the aggregation were head-to-tail and not head-to-head or tail-to-tail [29]. While this is problematic for high concentration expression and purification procedures,

the protein is relatively stable against aggregation at concentrations of 0.5 mg/ml and lower. At these concentrations, some characterization experiments have been carried out.

Our aim is to elucidate the structure of the $\alpha 1$ (XI) splice variants and their potential interactions with components of the extracellular matrix. This research will clarify the roles of collagen XI $\alpha 1$ chain isoforms in development and the regulated process of collagen fibrillogenesis. We report the isolation of $\alpha 1$ (XI) NTD[p7] and $\alpha 1$ (XI) NTD[p6b+7] using a recombinant expression system in *E. coli*. Our purification and refolding protocol produced an 11–12% yield of refolded protein. We have thus demonstrated an efficient method to produce enough refolded protein to perform future experiments to further characterize structural details and to determine the interactions of $\alpha 1$ (XI) NTD isoforms, which may be utilized for studying the mechanism that $\alpha 1$ (XI) uses to regulate fibrillogenesis in developing tissues.

References

1. Eyre D. Collagen of articular cartilage. *Arthritis Res* 2002;4:30–35. [PubMed: 11879535]
2. Keene DR, Oxford JT, Morris NP. Ultrastructural localization of collagen types II, IX, and XI in the growth plate of human rib and fetal bovine epiphyseal cartilage: type XI collagen is restricted to thin fibrils. *J Histochem Cytochem* 1995;43:967–979. [PubMed: 7560887]
3. Li Y, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, Oxford JT, Morris NP, Andrikopoulos K, Ramirez F, et al. A fibrillar collagen gene, *Col11a1*, is essential for skeletal morphogenesis. *Cell* 1995;80:423–430. [PubMed: 7859283]
4. Blaschke UK, Eikenberry EF, Hulmes DJ, Galla HJ, Bruckner P. Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. *J Biol Chem* 2000;275:10370–10378. [PubMed: 10744725]
5. Morris NP, Bachinger HP. Type XI collagen is a heterotrimer with the composition (1 α , 2 α , 3 α) retaining non-triple-helical domains. *J Biol Chem* 1987;262:11345–11350. [PubMed: 3112157]
6. Smith GN Jr, Williams JM, Brandt KD. Interaction of proteoglycans with the pericellular (1 α , 2 α , 3 α) collagens of cartilage. *J Biol Chem* 1985;260:10761–10767. [PubMed: 4030769]
7. Thom JR, Morris NP. Biosynthesis and proteolytic processing of type XI collagen in embryonic chick sterna. *J Biol Chem* 1991;266:7262–7269. [PubMed: 2016327]
8. Rousseau JC, Farjanel J, Boutillon MM, Hartmann DJ, van der Rest M, Moradi-Ameli M. Processing of type XI collagen. Determination of the matrix forms of the $\alpha 1$ (XI) chain. *J Biol Chem* 1996;271:23743–23748. [PubMed: 8798599]
9. Medeck RJ, Sosa S, Morris N, Oxford JT. BMP-1-mediated proteolytic processing of alternatively spliced isoforms of collagen type XI. *Biochem J* 2003;376:361–368. [PubMed: 12962540]
10. Zhidkova NI, Justice SK, Mayne R. Alternative mRNA Processing Occurs in the Variable Region of the Pro- $\alpha 1$ (XI) and Pro- $\alpha 2$ (XI) Collagen Chains. *J Biol Chem* 1995;270:9486–9493. [PubMed: 7721876]
11. Davies GB, Oxford JT, Hausafus LC, Smoody BF, Morris NP. Temporal and spatial expression of alternative splice-forms of the $\alpha 1$ (XI) collagen gene in fetal rat cartilage. *Dev Dyn* 1998;213:12–26. [PubMed: 9733097]
12. Morris NP, Oxford JT, Davies GB, Smoody BF, Keene DR. Developmentally regulated alternative splicing of the $\alpha 1$ (XI) collagen chain: spatial and temporal segregation of isoforms in the cartilage of fetal rat long bones. *J Histochem Cytochem* 2000;48:725–741. [PubMed: 10820146]
13. Oxford JT, Doege KJ, Morris NP. Alternative exon splicing within the amino-terminal nontriple-helical domain of the rat pro- $\alpha 1$ (XI) collagen chain generates multiple forms of the mRNA transcript which exhibit tissue-dependent variation. *J Biol Chem* 1995;270:9478–9485. [PubMed: 7721875]
14. Gregory KE, Oxford JT, Chen Y, Gambee JE, Gygi SP, Aebersold R, Neame PJ, Mechling DE, Bachinger HP, Morris NP. Structural organization of distinct domains within the non-collagenous N-terminal region of collagen type XI. *J Biol Chem* 2000;275:11498–11506. [PubMed: 10753969]

15. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, MR.; Appel, RD.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker, JM., editor. The Proteomics Protocols Handbook. Humana Press; Totowa: 2005. p. 571-607.
16. Lees JG, Smith BR, Wien F, Miles AJ, Wallace BA. CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal Biochem* 2004;332:285–289. [PubMed: 15325297]
17. Savitsky A, Golay MJE. Smoothing and differential of data by simplified least squares procedures. *Anal Chem* 1964;36:1627–1639.
18. Lobley A, Whitmore L, Wallace BA. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 2002;18:211–212. [PubMed: 11836237]
19. Whitmore L, Wallace BA. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* 2004;32:W668–673. [PubMed: 15215473]
20. Provencher SW, Glockner J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 1981;20:33–37. [PubMed: 7470476]
21. van Stokkum IH, Spoelder HJ, Bloemendal M, van Grondelle R, Groen FC. Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal Biochem* 1990;191:110–118. [PubMed: 2077933]
22. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 2000;287:252–260. [PubMed: 11112271]
23. Mao D, Wachter E, Wallace BA. Refolding of the H⁺-ATPase Proteolipid in Phospholipid Vesicles. *Biochemistry* 1982;21:4960–4968. [PubMed: 6291595]
24. Baynes BM, Wang DI, Trout BL. Role of arginine in the stabilization of proteins against aggregation. *Biochemistry* 2005;44:4919–4925. [PubMed: 15779919]
25. Moradi-Ameli M, de Chassey B, Farjanel J, van der Rest M. Different splice variants of cartilage alpha1(XI) collagen chain undergo uniform amino-terminal processing. *Matrix Biol* 1998;17:393–396. [PubMed: 9822205]
26. Seegmiller R, Fraser FC, Sheldon H. A new chondrodystrophic mutant in mice. Electron microscopy of normal and abnormal chondrogenesis. *J Cell Biol* 1971;48:580–593. [PubMed: 4100752]
27. Fallahi A, Kroll B, Warner LR, Oxford RJ, Irwin KM, Mercer LM, Shadle SE, Oxford JT. Structural model of the amino propeptide of collagen XI alpha1 chain with similarity to the LNS domains. *Protein Sci* 2005;14:1526–1537. [PubMed: 15930001]
28. Tsumaki N, Kimura T. Differential expression of an acidic domain in the amino-terminal propeptide of mouse pro-alpha 2(XI) collagen by complex alternative splicing. *J Biol Chem* 1995;270:2372–2378. [PubMed: 7836472]
29. Oxford JT, DeScala J, Morris N, Gregory K, Medeck R, Irwin K, Oxford R, Brown R, Mercer L, Cusack S. Interaction between amino propeptides of type XI procollagen alpha1 chains. *J Biol Chem* 2004;279:10939–10945. [PubMed: 14699108]

Fig. 1.

Schematic representation of the collagen type XI $\alpha 1$ chain. (A) The collagen type XI $\alpha 1$ chain is composed of an amino propeptide (Npp), a variable region (VR), a minor helix (mh), a telopeptide (tp), and a major helix (MH). The amino terminal domain (NTD) is composed of the Npp and VR. The Npp is made up of exons 1 through 5 and the variable region contains exons 6a, 6b, 7, 8 and 9. Exon 5 encodes a bone morphogenic protein 1 (BMP-1) processing site at AAQA/QHIDE, seven amino acids from the carboxyl terminus of the polypeptide region encoded by exon 5, indicated by arrow in A. (B) NTD[p7] contains exons 1 through 5, 7 and 9, is 289 amino acids in length and has a theoretical pI of 5.75. NTD[p6b+7] contains exons 1 through 5, 7, 6b and 9, is 340 amino acids in length and has a theoretical pI of 9.19.

Fig. 2.

Expression and extraction of recombinant proteins. Relative molecular mass markers (M_r) are shown in lane 1 of each gel. Pre (lane 2) and post (lane 3) IPTG induction of pET11a-NTD [p7] (A) and pET11a-NTD[p6b+7] (B). Arrows indicate the product NTD[p7] at approximately 45 kDa (A) and NTD[p6b+7] at approximately 50 kDa (B) relative molecular mass. Five sequential extractions of detergent-soluble proteins (lanes 2,4,6,8 and 10) and collection of inclusion bodies containing insoluble *E. coli* proteins and over expressed $\alpha 1(XI)$ NTD[p7] (C, lanes 3,5,7 and 9) and $\alpha 1(XI)$ NTD[p6b+7] (D, lanes 3,5,7 and 9). Material shown in lanes 9 of C and D were used for further purification.

Fig. 3.

Purification of recombinant proteins. (A) Representative chromatogram depicting purification of recombinant $\alpha 1(XI)$ NTD[p7] and the corresponding SDS-PAGE gel of fractions 10 – 23. (B) Representative chromatogram depicting purification of recombinant $\alpha 1(XI)$ NTD[p6b+7] and the corresponding SDS-PAGE gel of fractions 10 – 23. Relative molecular mass markers (M_r) are shown in lane 1 of each gel.

Fig. 4. Refolding of recombinant proteins. Relative molecular mass markers (M_r) are shown in lane 1 of each gel. (A) SDS-PAGE gel showing refolded (lane 2), dialyzed (lane 3) and concentrated (lane 4) $\alpha 1(XI)$ NTD[p7]. (B) SDS-PAGE gel showing refolded (lane 2), dialyzed (lane 3) and concentrated (lane 4) $\alpha 1(XI)$ NTD[p6b+7].

Fig. 5. Secondary structure analysis with circular dichroism. Circular dichroism spectra of $\alpha 1(XI)$ NTD[p7] (solid) and $\alpha 1(XI)$ NTD[p6b+7] (dashed) after refolding. Zero line is indicated for visual reference.

Table 1
Yield table for purification, denaturant removal and dialysis of $\alpha 1(XI)$ NTD[p7] and $\alpha 1(XI)$ NTD[p6b+7]
Yield was quantified by A_{280} and the Bradford assay. Molar extinction coefficients were calculated using ProtParam.

Step	$\alpha 1(XI)$ NTD[p7]				$\alpha 1(XI)$ NTD[p6b7]			
	Total protein ^a (mg)	Stepwise Yield (%)	Overall Yield (%)	Total protein ^a (mg)	Stepwise Yield (%)	Overall Yield (%)	Total protein ^a (mg)	Overall Yield (%)
Solubilization in 6 M GuHCl	1.693	100%	100%	0.380	100%	100%		
Purification	0.250	15%	15%	0.126	33%	33%		
Denaturant removal	0.192	77%	11%	0.115	91%	30%		
Dialysis	0.192	100%	11%	0.047	41%	12%		
Total weight per 100 mL culture	4.822			0.986				

^a Starting material was washed inclusion bodies from a 100 mL culture

Table 2
Table summarizing the secondary structures of NTD[p7] and NTD[p6+7]
Numerical values represent the relative content of each type of secondary structure as a percent of the total protein. Normalized root-mean-square deviation (NRMSE) is reported to indicate the goodness of fit of calculated data to experimental data. Npp data is included for comparison.

Isoform	α helix	β sheet	β turn	unordered	NRMSE
Npp	10%	33%	23%	34%	nd
NTD[p7]	17%	23%	25%	35%	0.165
NTD[p6b+7]	20%	26%	24%	30%	0.089