

# The $\alpha 1$ Chain of Collagen XI, Posttranslational Modifications, and the Pericellular Matrix

Katey Irwin: McNair Scholar

Dr. Julia Oxford: Mentor

Dept. of Biology



## Abstract

*Collagen type XI is a quantitatively minor but developmentally essential component of the extracellular matrix and as a secreted protein is subject to a variety of posttranslational modifications. These modifications have been shown to influence the affinity of endogenous binding between collagen XI and various extracellular matrix components. The three dimensional structure of the amino terminal domain (Npp) of this molecule has been determined through homology modeling. As a result a putative heparan sulfate binding site has been predicted. Heparan sulfate is an abundant sugar found on the surface of most cells and on proteoglycan molecules of the pericellular matrix. This study focuses on the structure of amino terminal domain of  $\alpha 1$  of collagen type XI and the subsequent interactions between this domain and other extracellular matrix proteins.*

## Introduction

In what many researchers are calling today's, "Post-Genomic" era, there has developed a great need for data analysis, management and facilitating decisions, to maintain dependable sample handling and adhere to protocols. Bioinformatics is the application of computer science to the interpretation and management of this biological data. The data generated from high throughput machines like the mass spectrometer, is enormous in size, and essential to the understanding of the new material being gathered every day. Central to protein analysis, is the gathering and understanding of the protein sequence. In recent years, LC-MS/MS has become an important tool in protein analysis and characterization. To better understand the structure and function of proteins, we must first understand what that sequence of the peptide is and what its posttranslational modification (PTM) are. These modifications are important in understanding protein physical and chemical properties, folding, stability, activity and function.

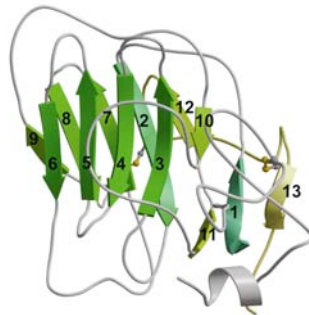


Figure 1. Schematic. Predicted model of the Npp domain of the  $\alpha 1$  chain of collagen XI molecule.



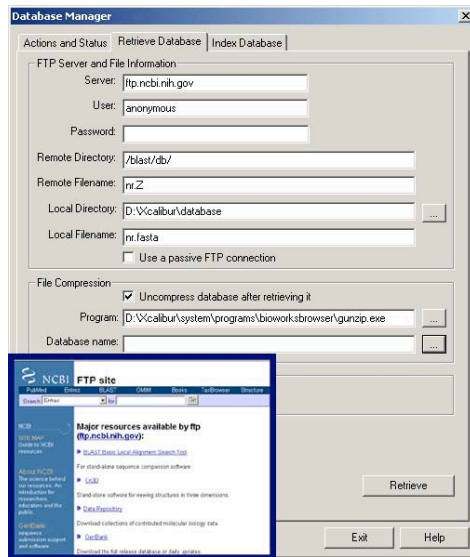


Figure 3. Downloading Databases. Databases are downloaded using BioWorks browser's Database manager.

The database used here was the non-redundant database or nr.fasta database. Once downloaded databases are then indexed to speedup search time and accuracy (Fig 4). Indexing involves pre-calculating the masses of the peptide sequences, based on specific enzyme digests, for each protein or nucleotide entry in the original *fasta* database, and storing this information in a new file.

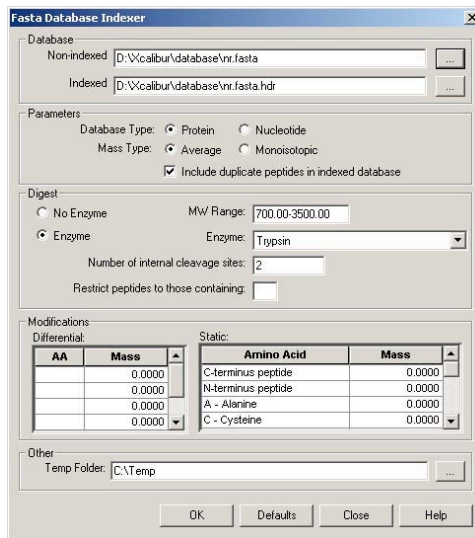


Figure 4. Indexing. Using BioWorks Indexer, databases can be modified to specific parameters to enhance search time.

## Sequest

An algorithm for performing protein identification & peptide sequencing. SEQUEST performs automated peptide/protein sequencing via database searching of MS/MS spectra without the need for any manual sequence interpretation, though it can make use of interpreted sequence information if available.

## Merlin

Merlin is Xcalibur's custom reports application. Customized reports are created containing all the results and important information produced by the mass spec analysis. Reports can contain textual summaries, as well as tabulated and graphical information.

## Results

Type XI collagen is a minor but essential component of the cartilage extracellular matrix. Alternative splicing encodes six alternatively expressed isoforms which add increasing structural and functional diversity to the amino-terminal of the pro- $\alpha 1(\text{XI})$  collagen gene (Fig 5).

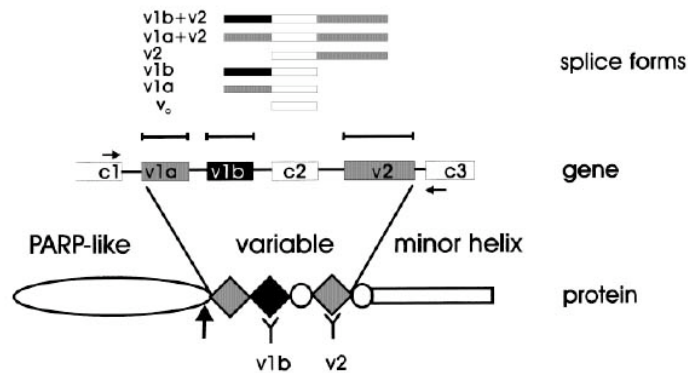


Figure 5. Schematic. The collagen type XI  $\alpha 1$  chain structure. Alternative splicing occurs generating six possible isoforms, as shown above.

This study focuses on determining how each isoform is posttranslationally modified. The variable region of the amino-terminal domain of  $\alpha 1$  of collagen XI isoforms was cloned into a recombinant vector, expressed in *E. coli* and isolated, purified. As the protein could not have posttranslational modification, the recombinant protein was used for control of MS and database analysis. Recombinant protein samples were separated by SDS-PAGE, excised and digested with trypsin. Samples were introduced into the mass spectrometer for analysis against a recombinant isoform-specific database. Database analysis allowed us to determine a set method to generate a high degree of sequence coverage (Fig 6).

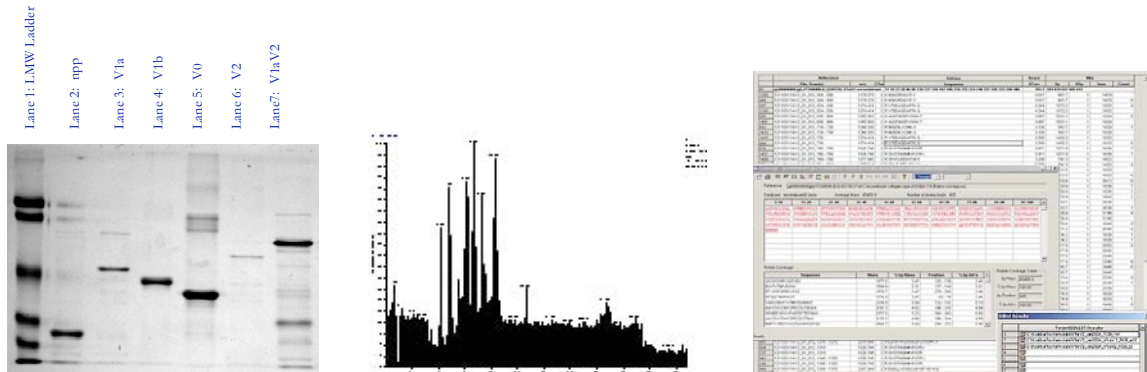


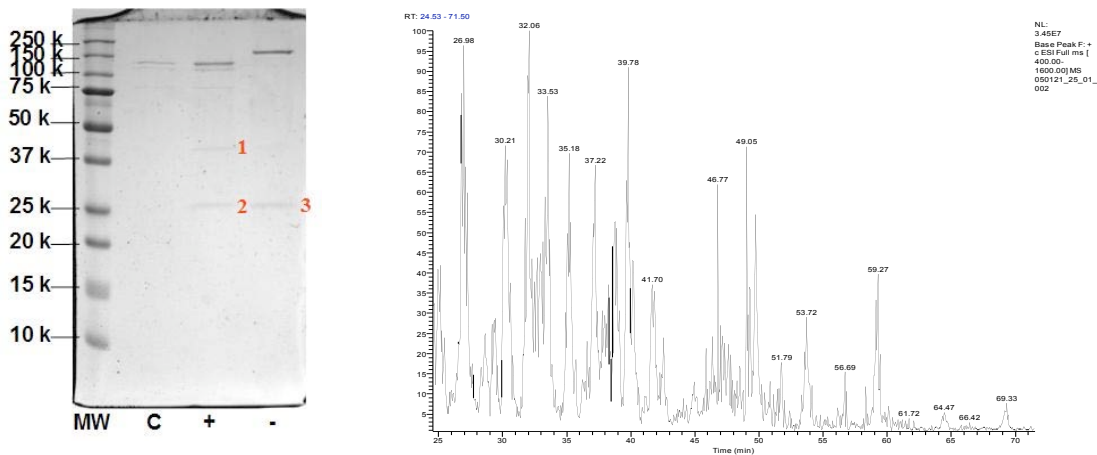
Figure 6. 1-D Analysis. A.) Coomassie stained SDS-PAGE gel of recombinant isoforms. V1aV2 (circled) was excised, digested and analyzed by LC-MS\MS. B.) V1aV2 chromatogram. C.) Data generated as result of three separate indexed searches of recombinant database resulted in nearly 100% coverage.

Results show obtained sequence coverage to be between 95% and 100% (Table 1). The sequence coverage information generated can then be used to assist in identification of posttranslational modifications.

Sequence Coverage					
Recombinant Isoform	V1a	V1b	VO	V2	V1aV2
Modified Trypsin	98.75	97.59	97.86	97.81	99.01
No enzyme	100	100	100	100	100

Table 1. Results of SEQUEST database searches using BioWorks. Sequence coverage is between 97-100%. Best results were obtained with no enzyme parameters set in BioWorks.

Studies are being conducted to identify the major protein components interacted with collagenous components in vivo. Homogenized cartilage was suspended in Tris buffer containing Guanidine hydrochloride and then centrifuged. The resulting pellet was then treated with collagenase. The 3 bands shown in Figure 7 were analyzed by mass spectrometer. The band 1 was identified as cartilage link protein 1. The band 2 and 3 were N-terminal domain of a1 of collagen XI.



Reference	Sequence	MH+	Charge	XC	Score	Accession	Sp	RSp	Peptides (Hits)
cartilage linking protein 1 [Bos taurus]					80.20	27805853.0			8 (8 0 0 0 0)
542	R.LLVEAEQAK.V	1001.16	2	2.49	0.45	589.3	1		13/16
546	R.LLVEAEQAK.V	1001.16	2	2.70	0.45	779.9	1		15/16
690 - 692	R.FVGFPDKK.H	938.11	2	2.19	0.25	745.7	1		13/14
826	R.NYGFWDKDK.S	1173.26	2	2.32	0.43	764.5	1		12/16
1038	R.FYYLHPTK.L	1182.40	2	2.18	0.59	650.0	1		12/16
1222 - 1224	K.VGQIFAAWK.L	1020.21	2	3.16	0.44	1342.3	1		15/16
1570	K.GGSDNDASLVITDLTLEDYGR.Y	2212.31	2	3.97	0.70	934.7	1		21/40
1574	K.GGSDNDASLVITDLTLEDYGR.Y	2212.31	2	3.54	0.67	910.3	1		19/40

Figure 7. Proteomics of Cartilage Sample. SDS-PAGE of 4M GuHCl insoluble component of cartilage treated with (+) or without (-) collagenase. C shows collagenase. MS spectra and the data analysis of band 1. Table shows database hit from band 1.

## **Discussion**

Research presented in this poster reflects the importance of identifying posttranslational modifications of collagen XI in vivo. This research has successfully identified the sequence of the six separate isoforms of collagen XI using mass spectrometry. We will apply this method to identify posttranslational modifications of collagen XI in cartilage. Further characterization of the posttranslational modifications may lead to a better understanding of the impact collagen XI on its surrounding pericellular matrix.

Additional research presented suggests that the cartilage link protein 1 may interact with collagenous components indirectly or directly. Further studies include additional analysis and characterization of the interaction between collagen and this identified cartilage link protein 1 both in vivo and in vitro.

The proteomics approach demonstrated in this study is becoming essential to research practices, providing the necessary databases to better understand the structure and function of proteins.

## **Acknowledgements**

This work was supported by grants from the Arthritis Foundation, NIH/NIAMS (RO1AR47985 and KO2AR48672) and by a grant from NIH/NCRR (P20RR16454) and the Boise State McNair Program. The author would like to thank Noriko Haziki-Taylor and Julia Thom Oxford.

## **References**

Fallahi, A., Kroll, B., Warner, L., Irwin, K., Mercer, L., Oxford, JRT. Forthcoming. "Structural model of the amino propeptide of collagen XI alpha 1 chain", *Protein Science*.