Short Note

MATRIX BIOLOGY

# Identification and Characterization of a Heparin Binding Site within the NC1 Domain of Chicken Collagen XIV

CLAIRE GIRY-LOZINGUEZ*, ELISABETH AUBERT-FOUCHER*, FRANÇOIS PENIN*, GILBERT DELÉAGE*, BERNARD DUBLET† and MICHEL VAN DER REST†

* Institute of Biology and Chemistry of Proteins, National Center for Scientific Research, Lyon and
† J. P. Ebel Institute of Structural Biology, Grenoble, France

## Abstract

Collagen XIV is known to bind to the dermatan sulfate chain of decorin and to the heparan sulfate chain of perlecan. To study its possible interaction with glycosaminoglycans, the NC1 domain of chicken collagen XIV was overproduced in *E. coli*. Purified NC1*(6–119)* appears poorly organized (the asterisks indicate the presence of extension sequences), but V8-protease generated fragments containing the 84–108 basic sequence tend to fold into α-helix. These fragments interact specifically with heparin, which induces an α-helical fold with a maximum effect for equimolar heparin/peptide ratio. These data demonstrate the existence of a glycosaminoglycan binding site in NC1.

Key words: collagen XIV, heparin binding site, non-collagenous domain.

## Introduction

Collagen XIV is a homotrimeric modular protein of the extracellular matrix which has been classified in the fibril-associated collagens with interrupted triple helix (FACIT) family (Shaw and Olsen, 1991). It contains two short triple helical domains (COL1 and COL2) and three non-collagenous domains (NC1, NC2 and NC3) (van der Rest and Bruckner, 1993).

Collagen XIV can interact with glycosaminoglycan (GAG) chains. It binds to the dermatan sulfate side chain of the small proteoglycan decorin (Font et al., 1993), to the heparan sulfate chains of the basement membrane proteoglycan perlecan and to heparin (Brown et al., 1993). The large NC3 domain of the molecule (obtained after collagenase digestion) or the COL1 and COL2 domains (isolated after pepsinization) fail to bind to collagen VI, to the heparan sulfate chain of perlecan (Brown et al., 1994) or to decorin (E. Aubert-Foucher, unpublished results).

The C-terminal NC1 domain of collagen XIV contains a RRKIAK sequence, which matches one of the two consensus sequences commonly found in sites for noncovalent binding of sulfated GAG chains (BBXB and BBBXXB, where B is a basic amino acid and X represents any amino acid) (Cardin and Weintraub, 1989). We overproduced this NC1 domain in *E. coli* in order to characterize its structure and GAG-binding properties.

Abbreviations used: CD, circular dichroism; FACIT, fibril-associated collagens with interrupted triple helix; GAG, glycosaminoglycan; NOE, nuclear Overhauser enhancement.; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

## Results and Discussion

Plasmid pNC1(6xHis), which is a derivative of p T7–7 (Cortay et al., 1994), constructed and expressed according to standard procedures (Sambrook et al., 1989), encoded the NC1 domain of chicken collagen XIV under T7 promoter control. The predicted primary sequence of the overproduced protein is:

6

**ARIRARGS**AGYGMGGGYGEPTDQDIPVVQLPHNS-YQIYDPEDLYDGEQQPYVV

HGSYPLPSPYSQSSYPSPHLAQPEFTPVREEMEAVEL-RSPGISRFRRKIAKRSIKTL

119

EHKRENAKEPSQ**LQHHHHHH**

The numbering refers to the cDNA derived NC1 sequence starting from the end of COL1 (residues 1 to 119 in Wächli et al., 1993). The N- or C-terminal extensions in bold were added for cloning and purification purposes. The expressed protein containing residues 6 to 119 of the NC1 sequence was designated NC1* (6–119)*. The asterisks indicate the presence of extension sequences.

Purification of NC1*(6–119)* by affinity chromatography on immobilized heparin was very efficient (see the legend of Fig. 1), and pure NC1*(6–119)* could be obtained after a second affinity chromatography on $Ni^{2+}$-nitriloacetic acid-agarose (lane 1, Fig. 1A). NC1* (6–119)* migrates in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with an apparent molecular mass of 21.4 kDa (14.8 kDa expected), but its genuineness was confirmed by N-terminal sequencing and mass spectrometry.

To localize the region of NC1*(6–119)* responsible for its interaction with heparin, this domain was digested with V8-protease, and the resulting polypeptides were tested for their ability to bind to heparin-Sepharose and/or to $Ni^{2+}$-nitriloacetic acid-agarose (Fig. 1A). After digestion, three bands can be resolved by SDS-PAGE (bands H, M and L, lane 2). Bands M and L interact with heparin (lane 4), but only the M band binds to $Ni^{2+}$-nitriloacetic acid-agarose (lane 6), indicating that it contains the 6xHis sequence. The slower migrating band (H), which doesn't bind to heparin, corresponds to the N-terminal part of NC1*(6–119)*. The M and L bands contain the heparin-binding sequence and correspond to the C-terminal end with and without the 6xHis extension, respectively. N-terminal sequencing of band M gives a triple sequence: MEAVELRSP, AVELRSPGI and LRSPGISRF. Thus, band M contains three polypeptides: NC1(82-119)*, NC1(84-119)* and NC1(87-119)* (Fig. 1C). The peptides of each band were separated by reverse phase high performance liquid chromatography as illustrated for the compounds of L band in Figure 1B. As expected, amino acid analysis and mass spectrometry revealed that the three peaks correspond to NC1(82-108), NC1(84-108) and NC1(87-108). A similar profile was obtained for band M, leading to the preparation of NC1(82-119)*, NC1(84-119)* and NC1(87-119)* fragments. Figure 1C summarizes the sequences of NC1-derived polypeptides which interact with heparin.

The circular dichroism (CD) spectra revealed that NC1*(6–119)* is poorly structured (Fig. 2A). Proton nuclear magnetic resonance analyses showed (i) a relatively low number of nuclear Overhauser enhancement (NOE) cross-peaks, (ii) the presence of broad peaks indicating conformational averaging processes and (iii) some NOE correlations typical of helical structures (not shown). Thus, NC1*(6–119)* should be considered as a mixture of interconverting conformers, exhibiting some local more stable folded regions. Moreover, NC1*(6–119)* showed a large propensity to fold upon stabilization with moderate amounts of trifluoroethanol (< 40%), indicating that NC1 could be folded in particular environments.

The secondary structure of basic 78–88 and 93–113 segments of NC1 are predicted as α-helices by most prediction methods (according to Geourjon and Deléage, 1995; data not shown). This is in agreement with the CD experiments showing that NC1(82–119)* and NC1(84–108) fragments exhibit a peak in the 220–230 nm range (Fig. 2A), indicating the presence of some helical folding. Moreover, these fragments showed a high propensity for helical folding on stabilization with trifluoroethanol. Of particular interest, addition of stoichiometric amounts of heparin increases the structuration of these fragments, as illustrated in Figure 2B for NC1 (84–108). The differential CD spectra with and without heparin showed the characteristic shape of helical peptides. These data demonstrate that heparin induces a specific conformational change of the basic NC1 fragments towards helix formation, suggesting that this region may be helical *in situ*.

Besides its helix inducing effect, heparin forms a 1:1 complex with NC1 fragments with an apparent dissociation constant lower than 2 μM in the presence of 0.15 M NaCl at pH 7.4 (Fig. 2C). The NaCl concentration must be raised to 500 mM to reverse heparin binding (Fig. 2D). These features are typical of well characterized heparin-binding sites (Cardin et al., 1991). Simi-

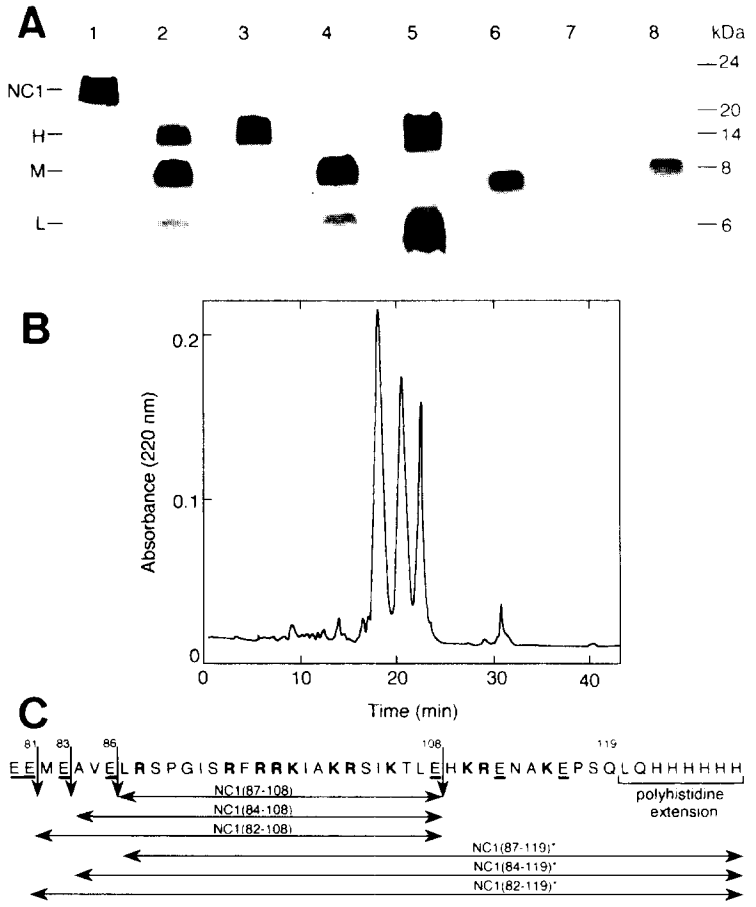


Figure 1. Identification of the heparin-binding NC1 fragments. After sonication in 10 mM Tris HCl at pH 7.45 (buffer A), the soluble bacterial extract was incubated in the presence of protease inhibitors with heparin-Sepharose. After extensive washing of the gel with 0.15 M NaCl in buffer A, NC1*(6–119)* was eluted with 0.5 M NaCl in the same buffer. The eluate was incubated with $Ni^{2+}$-nitriloacetic acid-agarose (Qiagen) and, after washing of the gel with 0.5 M NaCl and 0.1 M NaCl containing buffer A, the NC1*(6–119)* domain was finally eluted with 0.2 M imidazole, 0.1 M NaCl, pH 7.45. (A) SDS-PAGE analysis of purified NC1*(6–119)* and V8 protease-generated NC1*(6–119)* fragments. Pure NC1*(6–119)* (lane 1) was incubated with V8-protease (1/200, w/w) for 1 h 30 at 30 °C in 20 mM Tris HCl, 20 mM imidazole, 0.14 M NaCl, pH 7.45. An aliquot was withdrawn for electrophoretic analysis (lane 2). Two other aliquots (110 µg) were incubated in parallel for 45 min at 4 °C with 150 µl of heparin-Sepharose (lanes 3 and 4) or $Ni^{2+}$-nitriloacetic acid-agarose (lanes 5 and 6). Unbound fractions are shown in lanes 3 and 5. After washing of the gels with ten volumes of 10 mM Tris HCl, 0.15 M NaCl, pH 7.45, the material bound to heparin-Sepharose or $Ni^{2+}$-nitriloacetic acid-agarose was eluted with 10 mM Tris HCl, 0.75 M NaCl, pH 7.45 or 0.2 M imidazole, 0.1 M NaCl, pH 7.45 respectively (lanes 4 and 6). In lanes 7 and 8, a V8-protease NC1*(6–119)* digest was first bound to heparin-Sepharose, and the 0.75 M NaCl eluate was applied to a $Ni^{2+}$-nitriloacetic acid-agarose column. The unbound fraction is shown in lane 7 and the 0.2 M imidazole, 0.1 M NaCl, pH 7.45, eluate is shown in lane 8. All samples were analyzed by SDS-PAGE using 10% acrylamide gels, according to Schägger and Von Jagow (1987), and Coomassie Blue staining. H = higher band, M = medium band and L = lower band. (B) High performance liquid chromatography purification of material corresponding to lane 7 of Figure 1A. NC1-derived heparin-binding polypeptides were separated on an Aquapore C8 RP300 column (Brownlee) using an aqueous acetonitrile gradient (4–40%) over 45 min in 0.1% trifluoroacetic acid. A similar profile is obtained with material corresponding to lane 8 of Figure 1A (not shown). (C) Sequences of NC1*(6–119)* heparin-binding fragments obtained. The vertical arrows represent V8-protease cleavage sites.

lar results were obtained for the other basic NC1 fragments, even for the shorter one (NC1(84-108)). All these data indicate that this region of NC1 is able to bind reversibly to heparin or, more probably, to specific gag molecules. The possible location of the heparin-binding site in NC1(XIV) was investigated in accord with the spacial α-helix heparin-binding pattern described by Margalit et al. (1993). Due to the numerous basic residues, as many as four potential motifs may satisfy Margalit's criteria: NC1(97–110), NC1(98–111), NC1(88–101) and

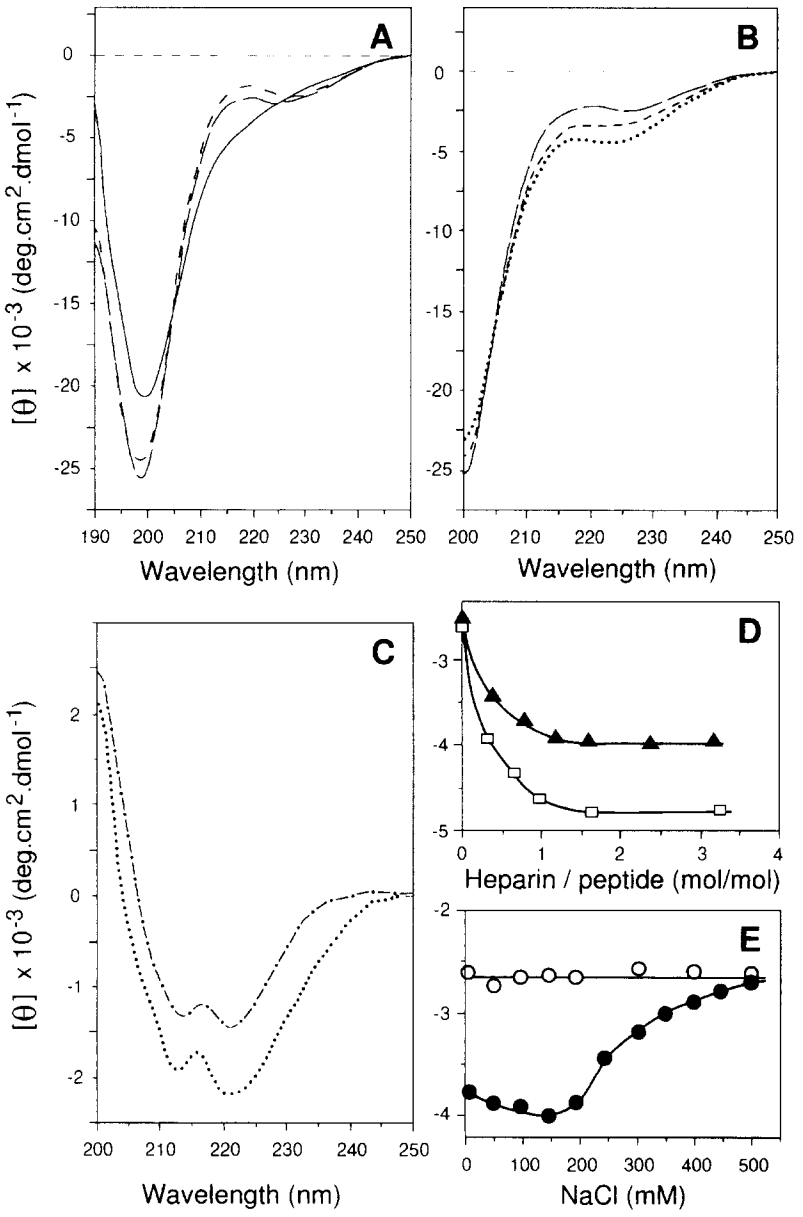


Figure 2. Circular dichroism (CD) spectra of NC1 and NC1 fragments and specificity of heparin binding. All far-UV CD spectra were recorded at room temperature with a Jobin-Yvon CD6 in 10 mM sodium phosphate, pH 7.5, in the absence (A) or in the presence (B, C and D) of 150 mM NaCl at a protein concentration of 15 μM in dichroically neutral quartz cuvettes with an optical pathlength of 2 mm. Samples were scanned in 0.2 nm increments, integration time 2 s. Baseline was established with buffer alone and baseline-corrected spectra were smoothed by using a third-order least squared polynomial fit. When heparin (Sigma, 6 kDa average mass) was present, the spectra were corrected for its contribution (note that at the concentrations used, heparin contribution to the CD is negligible). Because of the noise induced by the presence of chloride ions, CD spectra could not be recorded under 200 nm. All calculations were performed using the CD6 spectrum processing software. Peptide concentration was determined by amino acid analysis. Data were expressed in mean residue molar ellipticity. (A) Spectra of NC1(6–119)* (solid line), NC1 (82–119)* (small dash line) and NC1(84–108) (long dash line), respectively. (B) CD spectra of NC1(84–108) in the absence (long dash line) or presence of 0.5 (small dashed line) and 1 (dotted line) mol of heparin per mol of peptide. (C) Differential spectra between spectra recorded in the presence and absence of stoichiometric amounts of heparin for NC1(84–108) fragment (dotted line; this differential spectrum was calculated from the spectra obtained in (B) and for NC1(82–119)* fragment (dot-dash line). (D) Heparin binding on NC1(82–119)* and NC1(84–108) monitored by ellipticity at 222 nm (closed triangles and open squares, respectively). The NC1 fragments (15 μM) were mixed with heparin at the indicated ratio in 10 mM sodium phosphate, pH 7.5, containing 150 mM NaCl. (E) Effect of NaCl on the heparin binding to NC1(82–119)* monitored by ellipticity at 222 nm. The NC1(82–119)* fragment (15 μM) was incubated in the absence (open circle) or presence (closed circle) of an equimolar concentration of heparin in 10 mM sodium phosphate, pH 7.5, with the indicated amounts of NaCl before recording the CD spectra.

NC1 (102–115). Although the putative NC1(88–101) motif contains the heparin-binding-type sequence BBBXXB, none of the other motifs can be discarded *a priori*. Thus, the heparin binding site of NC1 appears more complicated than generally observed for the consensus heparin-binding sequence. This complexity may be related to the trimeric nature of NC1 in native collagen and is probably related to the recognition of specific GAG motifs.

As previously described (Font et al.,1993), immobilized collagen XIV binds decorin in solid phase assays. This binding is inhibited by very low amounts of heparin, demonstrating that this GAG can compete with the dermatan sulfate side chain of decorin. However, overproduced NC1*(6–119)* failed to bind decorin (not shown), indicating that its heparin-binding site cannot account by itself for the decorin-binding property of collagen XIV. This may reflect the necessity of a putative additional decorin binding site to stabilize the interaction with collagen XIV. This site may reside on the neighboring domains (COL1, NC2, COL2). Another possibility is that the trimerization of NC1 segments in intact collagen is required for binding on decorin.

## Acknowledgements

We thank E. Forest for mass spectrometry analysis, J.C. Cortay for its advises in construction of plasmids, A. Lesage for her help in NMR experiments, D. Gerecke and P. Castagnola for providing collagen XIV cDNA clones and the PS1 antiserum to chicken collagen XIV, L.C. Rosenberg for the generous gift of decorin, M.M. Boutillon for amino acid and N-terminal sequences analysis and A Bosch and C. Van Herrewege for their help in the preparation of illustrations. This work was supported by the Centre National de la Recherche Scientifique (CNRS).

## References

Brown, J.C., Golbik, R., Mann, K. and Timpl, R.: Structure and stability of the triple-helical domains of human collagen XIV. *Matrix Biol. 14:* 287–295, 1994.

Brown, J.C., Mann, K., Wiedemann, H. and Timpl, R.: Structure and binding properties of collagen type XIV isolated from human placenta. *J. Cell Biol. 120:* 557–567, 1993.

Cardin, A.D., Demeter, D.A., Weintraub, H.J.R. and Jackson, R.L.: Molecular design and modeling of protein-heparin interactions. *Meth. Enzymol. 203:* 556–583, 1991.

Cardin, A.D. and Weintraub, H.J.R.: Molecular modeling of protein-glycosaminoglycan interactions. *Arteriosclerosis 9:* 21–32, 1989.

Cortay, J.C., Nègre, D., Scarabel, M., Ramseier, T.M., Vartak, N.B., Reizer, J., Saier, Jr., M.H. and Cozzone, A.J.: *In vitro* asymmetric binding of the pleiotropic regulatory protein, FruR, to ace operator controlling glyoxylate shunt enzyme synthesis. *J. Biol. Chem. 269:* 14885–14891, 1994.

Font, B., Aubert Foucher, E., Goldschmidt, D., Eichenberger, D. and van der Rest, M.: Binding of collagen XIV with the dermatan sulfate side chain of decorin. *J. Biol. Chem. 268:* 25015–25018, 1993.

Geourjon, C. and Deléage, G.: A 3D module coupled to protein sequence analysis methods. *J. Mol. Graph. 13:* 209–212, 1995.

Margalit, S., Fischer, N. and Ben-Sasson, S.A.: Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues. *J. Biol. Chem. 268:* 19228–19231, 1993.

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

Shaw, L.M. and Olsen, B.R.: FACIT collagens: Diverse molecular bridges in extracellular matrices. *Trends Biochem. Sci. 16:* 191–194, 1991.

Sugrue, S.P., Gordon, M.K., Seyer, J., Dublet, B., van der Rest, M. and Olsen, B.R.: Immunoidentification of type XII Collagen in embryonic tissues. *J. Cell Biol. 109:* 939–945, 1989.

van der Rest, M. and Bruckner, P.: Collagens: Diversity at the molecular and supramolecular levels. *Curr. Opin. Struct. Biol. 3:* 430–436, 1993.

Wächli, C., Trueb, J., Kessler, B., Winterhalter, K.H. and Trueb, B.: Complete primary structure of chicken collagen XIV. *Eur. J. Biochem. 212:* 483–490, 1993.

Dr. Michel van der Rest, Institut de Biologie Structurale J.P. Ebel, 41 av. des Martyrs, 38027 Grenoble Cedex 01, France

Received September 15, 1997; accepted March 15, 1998