Structural Analysis of the Heparin-Binding Site of the NC1 Domain of Collagen XIV by CD and NMR[†,‡]

Roland Montserret,[§] Elisabeth Aubert-Foucher,[§] Michael J. McLeish,[||,⊥] Joanna M. Hill,[||] Damien Ficheux,[§] Michel Jaquinod,[#] Michel van der Rest,[#] Gilbert Deléage,[§] and François Penin*[,§]

Institut de Biologie et de Chimie des Protéines, CNRS UPR 412, 7 passage du Vercors, 69367 Lyon Cedex 07, France, Victorian College of Pharmacy, Monash University, 381 Royal Pde, Parkville 3052, Australia, and Institut de Biologie Structurale Jean Pierre Ebel, CEA-CNRS, 43 avenue des Martyrs, 38027 Grenoble Cedex 1, France

Received January 5, 1999; Revised Manuscript Received March 11, 1999

ABSTRACT: Type XIV collagen, a fibril-associated collagen with interrupted triple helices (FACIT), interacts with the surrounding extracellular matrix and/or with cells via its binding to glycosaminoglycans (GAGs). To further characterize such interactions in the NC1 domain of chicken collagen XIV, we identified amino acids essential for heparin binding by affinity chromatography analysis after proteolytic digestion of the synthetic peptide NC1(84−116). The 3D structure of this peptide was then obtained using circular dichroism and NMR. The NC1(84−116) peptide appeared poorly structured in water, but the stabilization of its conformation by the interaction with hydrophobic surfaces or by using cosolvents (TFE, SDS) revealed a high propensity to adopt an α-helical folding. A 3D structure model of NC1(84−116), calculated from NMR data recorded in a TFE/water mixture, showed that the NC1-heparin binding site forms a amphipathic α-helix exhibiting a twisted basic groove. It is structurally similar to the consensus spatial α-helix model of heparin-binding [Margalit et al. (1993) *J. Biol. Chem. 268*, 19228−19231], except that the GAG binding domain of NC1 may be extended over 18 residues, that is, the NC1(94−111) segment. In addition, the formation of a hydrophobic groove upon helix formation suggests the contribution of additional sequences to ensure the stability of the GAG-binding domain. Overall the NC1(84−116) model exhibits a nativelike conformation which presents suitably oriented residues for the interaction with a specific GAG.

Collagens are the most abundant proteins present in the extracellular matrix. They are characterized by a triple helical structure consisting of the repetitive sequence Gly-Xaa-Yaa (where Xaa is often a proline residue and Yaa is often a hydroxyproline residue). They have been classified according to their structural organization in the extracellular matrix (for reviews, see refs *1*−*3*). Among the nonfibrillar collagens, a class called FACITs[1] for fibril-associated collagens with interrupted triple helix (*4*, *5*) has been defined since their discovery in the neighborhood of quarter-staggered fibrils. In the FACIT's family, collagen IX is clearly associated at the surface of the fibers, and covalent links between collagen IX and the fibrils forming collagen II have been demonstrated (*6*, *7*). For other members of the family, like collagens XII and XIV, the mode of association has yet to be identified. However, it has been proposed that proteoglycans, such as decorin, could provide a link between collagen XIV and the fibrils through their glycosaminoglycan (GAG) chains (*8*).

Collagen XIV is a homotrimeric protein which contains two short triple helical domains (COL1 and COL2) and three non-triple helical domains (NC1, NC2, and NC3). Intensive study has been undertaken to understand the trimeric assembling of collagen XIV (*9*), its overall function, and the function of its different domains (*8*, *10*, *11*). As is the case for numerous proteins of the extracellular matrix, the interaction of collagen XIV with glycosaminoglycans (GAGs) is well-documented. Apart from its binding with the dermatan sulfate chain of decorin (*8*), it has also been found that collagen XIV interacts with the heparan sulfate chain of perlecan, with heparin (*10*), and with a proteoglycan form of the cellular receptor CD44 (*12*). Two sites involved in GAG binding have been described for collagen XIV and

[†] This work was supported by the Centre National de la Recherche Scientifique, by a Program Emergence grant from the Région Rhône-Alpes (France), and by the Monash Research Fund (Australia).

[‡] The coordinates have been deposited with the Brookhaven Protein Data Bank under the accession codes 1B9P (average structure), 1B9Q (set of 19 structures), and R1B9QMR (NMR restraints).

* To whom correspondence should be addressed. E-mail: f.penin@ibcp.fr. Fax: 334 72 76 90 50.

[§] Institut de Biologie et de Chimie des Protéines.

[||] Monash University.

[⊥] Present address: College of Pharmacy, University of Michigan, Ann Arbor MI 48109.

[#] Institut de Biologie Structurale.

[1] Abbreviations: 2D and 3D, two and three-dimensional; CD, circular dichroism; cmc, critical micellar concentration; DQF-COSY, double quantum filtered correlation spectroscopy; DTNB, 5,5′dithiobis-(2-nitrobenzoic acid); DTT, dithiothreitol; FACIT, fibril-associated collagen with interrupted triple helices; FID, free induction decay; GAG, glycosaminoglycan; HMQC, heteronuclear multiple quantum coherence; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; NMR, nuclear magnetic resonance; ppm, parts per million; PMSF, phenylmethylsulphonyl fluoride; rmsd, root-mean-squared deviation; ROESY, rotating frame Overhauser effect spectroscopy; RP-HPLC, reversed-phase high performance liquid chromatography; SCUBA, stimulated cross-peaks under bleached alphas; TFE, 2,2,2-trifluoroethanol; TOCSY, total correlation spectroscopy; TPPI, time proportional phase increment; TSP-*d*4, 3-(trimethysilyl)propionic-2,2,3,3-d_4 acid; WATERGATE, water suppression by gradient-tailored excitation.

10.1021/bi9900222 CCC: $18.00
Published on Web 04/29/1999

were located at the opposite extremities of the protein. The first site was assigned to the NH_2-terminal fibronectin type III repeat located in the NC3 domain of human collagen XIV and is thought to bind to decorin and to whole cells (*13*, *14*). The second site was recently demonstrated in the C-terminal NC1 domain of chicken collagen XIV by studies showing that recombinant NC1 bound strongly to heparin by both affinity chromatography (*11*) and solid-phase assays (E. Aubert-Foucher, unpublished experiments). A physiological GAG partner has not yet been found for this site. In addition, cDNA analyses have revealed the existence of variants of type XIV collagen which differ in NC1 and NC3 sequences (*15*−*17*) and it is tempting to speculate that different forms of collagen XIV exhibiting distinct GAG-binding properties are present in tissue. Further, it is now generally accepted that extracellular matrix proteins are proteolytically processed in various normal or pathologic physiological conditions and there are several examples reported showing that such domains may exhibit heparin-binding activity. We have shown that the NC1 domain of FACIT collagen XII (which is very similar to NC1 of collagen XIV) is cleaved in various cell media (Aubert-Foucher et al., submitted for publication), and it is possible that the C-terminal processing of collagen XIV in vivo may occur, resulting in the release of NC1 (or part of NC1). Such hypothetical release of NC1 in combination with the specific interaction with GAGs may be of major physiological importance.

Knowledge of the mechanisms by which collagens interact with GAGs such as heparin/heparan sulfate is of great importance when trying to understand the role of such interactions in the structure−function relationship of collagens, and in the modulation of cell and matrix interactions. Heparin and heparan sulfate are negatively charged polymers and are thought to interact primarily with proteins via electrostatic interaction. However, it is clear that the specific interaction of proteins with GAG depends not only on the linear spacing of basic residues on specific structural motifs (*18*, *19*) but also upon the structural folding, that is, the spacial relationship of distant domains and the quaternary structure (*20*−*23*). In this context, detailed experimentally derived structural information is mandatory for a better understanding of GAG−protein interaction. Unfortunately, only limited experimental information is available on the conformation of heparin-binding domains in collagens (*21*, *24*). While the complexity of collagen molecules is an issue, this is mainly due to the difficulties in obtaining large amounts of the purified, native, and homogeneous trimeric domains that are required for structural analysis. One approach to overcome these limitations and obtain insights at the molecular level for a better understanding of GAG binding is to carry out the conformational analysis of synthetic peptides containing GAG-binding activities.

We have recently located the NC1 heparin-binding site of chicken collagen XIV in the basic 84−108 region by analyzing NC1 protease digestion products for their ability to bind to heparin (*11*). Moreover, we showed by CD that the binding of heparin tends to induce some α-helical folding of NC1(84−108) and NC1(84−119) fragments which were poorly folded in water (*11*). To better characterize the structure of the GAG-binding site of NC1, we chemically synthesized the NC1(84−116) peptide. Heparin affinity chromatography of peptide fragments obtained after NC1-(84−116) digestion by various proteases allowed the identification of the essential residues involved in heparin binding. In addition, to assess NC1(84−116) conformational preference in situ, we analyzed its structure under various stabilizing conditions by using cosolvents (TFE, SDS) that mimic the environment found in the hydrophobic core of globular proteins (*25*−*33*). An alternative way to approximate the biologically relevant conformations of protein-bound peptide segments is to use hydrophobic surfaces, and recently, a CD procedure has been developed which allows the analysis of the conformation of peptides bound to C18 chains covalently linked to quartz plates (*34*, *35*). We show here that, whatever the method and the conditions used (i.e., interaction either with hydrophobic plates or with SDS, or by stabilization with TFE), NC1(84−116) exhibits a high propensity to fold into α-helix and it is likely that such folding may occur in a physiological context such as the interaction with GAG. To determine this precise conformation, we have used NMR spectroscopy to determine the 3D structure of NC1(84−116) in the presence of a water−TFE mixture. The molecular model we obtained shows the positioning of the main basic residues involved in heparin binding, and provides a framework for a molecular level understanding of the mechanism by which this region of collagen XIV interacts with GAG.

EXPERIMENTAL PROCEDURES

Peptide Synthesis. The peptide was chemically synthesized using the stepwise solid phase of Merrifield with a Fmoc (9-fluorenylmethyloxycarbonyl) chemistry in a Millipore 9050 plus continuous flow synthesizer. The C-terminus carboxylate group was synthesized in the amide form, and the N-terminus amino group was acetylated with acetic anhydride to give greater similarity to the parent protein. After cleavage and deprotection, the peptide was purified by preparative RP-HPLC on a Vydac C18 column using 0.1% aqueous trifluoroacetic acid and an acetonitrile gradient. To avoid intermolecular disulfide bond formation during CD and NMR studies, we converted the N-terminal cysteine residue into *S*-carboxyamidomethyl derivative using iodoacetamide.

NC1(84−116) Fragmentation by Proteolysis and Analysis of the Resulting Peptides by Affinity Chromatography on Immobilized Heparin. The NC1(84−116) peptide was first reduced with DTT, and all subsequent steps (blocking of the N-terminal cysteine and digestion with chymotrypsin or plasmin) were performed on the peptide bound to heparin-Sepharose as follows: NC1(84−116) in 0.15 M NaCl, 10 mM Hepes pH 7.4 (buffer A) was reduced with 20 mM DTT at 33 °C for 40 min and batch incubated with heparin-Sepharose CL6B (Pharmacia) for 20 min at room temperature. All the peptide was bound to the affinity matrix. After being extensively washed with buffer A, NC1(84−116) was treated four times with 1 mM DTNB (i.e., until the reacting buffer was colorless) and the gel washed again. Aliquots of the heparin gel bearing the resulting blocked peptide were batch incubated for 30 min at room temperature in the absence or in the presence of α-chymotrypsin or plasmin (2.5 μg/mL, Sigma) in 0.15 M NaCl, 10 mM Tris-HCl pH 8.0 buffer containing 1 mM $CaCl_2$. PMSF (0.5 mM) was added, and the unbound fractions were collected. After

extensive washing of the resin with buffer A containing PMSF, bound peptides were sequentially eluted with 0.5 and 1 M NaCl in 10 mM Tris-HCl, pH 7.4, and the eluates pooled to provide the bound fraction. All of the peptides were purified by reverse-phase HPLC on an aquapore C8 RP300 column (Brownlee) using aqueous acetonitrile gradients in the presence of 0.1% trifluoroacetic acid. Once purified, the peptides present in the bound fraction were tested again for their ability to interact with heparin-Sepharose by batch incubation as described above. Finally, the peptides were identified by MALDI mass spectroscopy as described previously (*36*) using a time-of-flight mass spectrometer (Voyager Elite Xl, Perseptive Biosystems, Framingham, MA) equipped with a 337 nm nitrogen laser.

Circular Dichroism Measurements. CD spectra were recorded either on a Jobin-Yvon CD6 or on a Jasco J-710 spectrometer equipped with variable temperature units and calibrated with ammonium D-10-camphorsulfonate. Routinely, measurements were carried out at 298 K using a 0.1 cm path length quartz cuvette (Hellma) with peptide samples at concentrations ranging from 20 to 110 μM in 10 mM sodium phosphate buffer, pH 7.4. Spectra were recorded in the 190−250 nm wavelength range with 0.2 nm increments and 2 s integration time. The baseline-corrected spectra were smoothed by using a third-order least-squares polynomial fit. Calculations were performed using either the CD6 or the Jasco processing softwares. Assuming that the residue molar ellipticity at 222 nm is exclusively due to α-helix, the α-helical content was estimated according to Chen et al. (*37*).

Quartz plates, coated with a C_4, C_8, or C_{18} hydrophobic surface, were prepared using a modification of the method of Blondelle et al. (*34*, *35)*. The plates were derivatized with octadecyldimethylchlorsilane (C_{18}), octyldimethylchlorosilane (C_8), or *n*-butyldimethyl-chlorosilane (C_4) in anhydrous toluene (*38*). NC1(84−116) was bound to a set of nine plates by incubating them overnight with a solution of peptide (50 μM) in 5 mM MOPS, pH 7.0, at room temperature. After the plates were rinsed with milli-Q water to remove unbound peptide, the nine plates were then fitted into a 2 cm quartz cell containing either 10 mM sodium phosphate buffer, pH 7.4, or milli-Q water. Spectra were acquired at 293 K with a scan speed of 20 nm/min over the range 250−190 nm in 0.2 nm increments, and with a 1 s integration time. The results obtained were an average of at least 4 scans. A spectrum of the plates without peptide bound provided the blank.

NMR Spectroscopy. Lyophilized peptide was dissolved in 10 mM sodium phosphate buffer at pH 6.0 containing 50% TFE-d_2 (2,2,2-trifluoroethyl-1,1-d_2 alcohol >99% isotopic enrichmen). All NMR experiments were recorded at 500 MHz on a Varian Unity-*plus* spectrometer equipped with *ultra-nmr* shims and using a triple resonance proton−carbon−nitrogen 5 mm probe with a self-shielded z gradient coil. Spectra were acquired nonspinning at temperatures ranging from 283 to 303 K. For 2D homonuclear ^{1}H experiments, DQF-COSY, clean-TOCSY, NOESY, and ROESY were performed according to the conventional pulse sequences (references in *9*, *39*). For NOESY, mixing times in the 35−250 ms range were used to evaluate spin diffusion. A spin-lock of 150 ms obtained by continuous irradiation was used for ROESY experiments. Water suppression was carried out using selective, low-power irradiation during the 1.5 s relaxation delay and during the mixing time in NOESY experiments. Alternatively, water was suppressed using a WATERGATE sequence before detection. Data collection and processing were carried out as detailed previously (*9*, *39*). For ^{13}C-^{1}H experiments, broad band ^{13}C decoupling during acquisition was accomplished by means of a GARP-1 sequence. gHSQC (gradient version of HSQC) was recorded with 2000 data points and 128 scans in t_2 dimension and 512 increments in t_1. A 50 ms spin lock was used for the gHSQC-TOCSY, with 2000 data points acquired in 256 scans in t_2 per 360 increments in t_1.

NMR-Derived Constraints. NOE intensities used as input for the structure calculations were obtained from the NOESY spectrum recorded with a 50 ms mixing time to avoid the spin diffusion arising at longer mixing times. Spectra obtained at 283 and 303 K were also used to estimate NOE intensities for cross-peaks unresolved at 293 K. Acquisition of a ROESY spectrum was always performed for each set of experiments to check that no NOESY cross-peak was missing. NOEs were partitioned into three categories of intensities that were converted into distances ranging from a common lower limit of 1.8 Å to upper limits of 2.6, 3.8, and 5 Å, respectively. The cross-peak intensity of the H^{δ}-H^{ϵ} protons of Phe95 was used as a distance reference (2.45 Å). Protons without stereospecific assignments were treated as pseudoatoms, and correction factors were added to the upper and lower distance constraints (*40*). NOEs back-calculations were performed from calculated structures by using the standard procedure of X-PLOR 3.1 (*41*).

Structure Calculations. Three-dimensional structures were generated from NOE distances with X-PLOR 3.1 program (*41*), using the standard force fields and default parameter sets, except some minor modifications to increase the duration of the molecular dynamic simulations and the number of energy minimization steps as detailed previously (*9*, *39*). Sets of 50 structures were calculated to widely sample the conformational space, and the structures were analyzed for NOE violations. The selected structures were compared by pairwise rmsd for the backbone atom coordinates (N, Cα, and C′). Local analogies were analyzed by calculating the local rmsd of a tripeptide window slided along the sequence. Statistical analysis, superimposition of structures, 3D graphic displays, and manipulations were achieved by using either ANTHEPROT 2.0 (*42*) or RASMOL 2.6 (R. Sayle program) software. The secondary structure elements and Ramachandran plots were analyzed according to the Kabsch-Sander definition rules, as incorporated in the program PROCHEK (*43*).

RESULTS

For experimental structural investigation of the previously identified NC1 heparin-binding region (*11*), a fragment was chemically synthesized with the sequence C*A_{84}VELRS-PGISRFRRKIAKRSIKTLEHKRENAKE$_{116}$ and denoted NC1(84−116). The numbering of NC1 refers to that proposed (*15*). Note that the additional cysteine residue in the N-terminal position (C*) was added for labeling or covalent-binding purposes. However, it was modified into *S*-thionitrobenzoate cysteine for heparin-binding studies or *S*-carboxamidomethyl cysteine for CD and NMR studies to avoid the potential formation of intermolecular disulfide bonds.

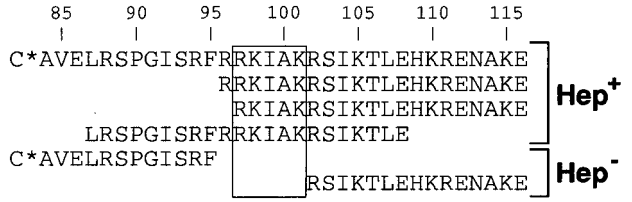

FIGURE 1: Localization of the amino acids of NC1(84−116) essential in heparin binding. The whole NC1(84−116) peptide and its proteolytic fragments were produced, purified, and tested for their ability to bind to heparin-Sepharose as described in the Experimental Procedures. Peptides R96-E116 and C*A84-F95 were produced after α-chymotrypsin digestion, whereas peptides R97-E116 and R102-E116 were obtained after NC1(84−116) proteolysis by plasmin. In addition, the peptide L87-E108 was included in this figure, as we have previously shown that it binds to heparin (*11*). Hep+ and Hep− refer to fragments that can or cannot bind to immobilized heparin, respectively. C*: *S*-thionitrobenzoate cysteine.

Identification of Residues Essential for Interaction with Heparin. To determine the minimal residues required for interaction with heparin, we bound NC1(84−116) to immobilized heparin before digestion with α-chymotrypsin or plasmin in order to protect the cleavable residues involved in the heparin-binding site. Bound and unbound peptides were further purified by reverse-phase HPLC, and the heparin-binding ability of each peptide was tested again. This strategy allowed the elimination of peptides potentially retained on the immobilized heparin medium because of their complexion with heparin-bound peptides. Figure 1 shows that only the peptides containing the RKIAKRSIKTLE sequence (97−108 segment) bind to heparin. In addition, the R102-E116 and C*A84-F95 peptides, which did not contain the RKIAK sequence (97−101 segment), were not retained on heparin, indicating the essential role of these residues in heparin binding. The lack of ability of the R102-E116 peptide to bind to heparin also indicates that residues R102 and K105 are likely to play a secondary role in heparin binding. These results also indicate that, in themselves, the basic cluster 109−111 (HKR) and K115 are not sufficient to promote binding to heparin. This is confirmed by the heparin-binding ability of the L87-E108 peptide which does not contain these residues. In summary, the RKIAK segment (97−101) appears to be absolutely required for heparin binding, but the minimal heparin-binding site possibly also involves at least the neighboring basic residues R102 and K105.

Circular Dichroism Analysis. Figure 2 shows the CD spectra of NC1(84−116) in the presence of various concentrations of TFE or SDS (A and B, respectively), or in interaction with hydrophobic surfaces (C). The CD spectrum of NC1(84−116) in 10 mM sodium phosphate buffer pH 7.4 (Figure 2A) shows an intense negative band at 198 nm characteristic of a peptide in a random-coil conformation. Modifying the pH of the buffer from 4.0 to 9.0 and/or increasing the ionic strength did not significantly modify the spectrum (data not shown).

In the presence of increasing concentrations of TFE (Figure 2A), the shape of the spectrum becomes increasingly typical of α-helical folding with two minima at 208 and 222 nm and a maximum at 192 nm. This process occurs while the TFE concentration goes from 0% to 41% (v/v) and is maximal at 41% TFE. Between 41% and 59% TFE (and up to 75%, data not shown), only minor changes in the

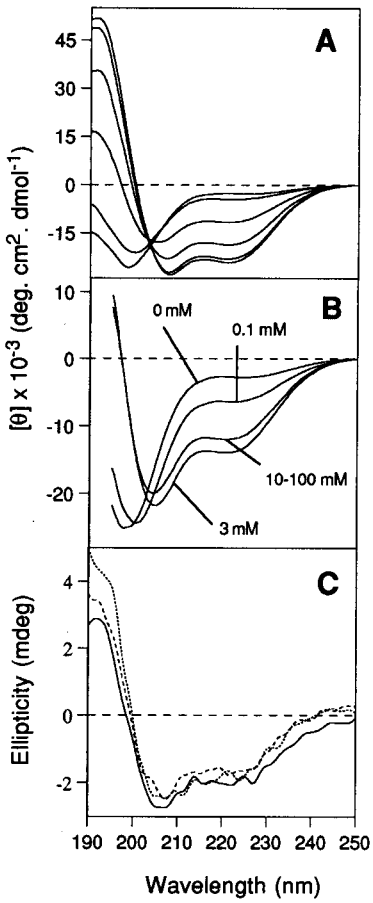

FIGURE 2: Far-UV CD spectra of NC1(84−116). (A) TFE titration of 15 μM NC1(84−116) in 10 mM sodium phosphate buffer, pH 7.4. The TFE concentrations on the spectra from top to bottom (at 222 nm) are the following: 0%, 10%, 18%, 25%, 41%, 59%. (B) Spectra of 15 μM NC1(84−116) in 10 mM sodium phosphate buffer, pH 7.4, in the presence of SDS. Note that all of the spectra recorded in the 10 to 100 mM SDS range are superimposable. (C) Spectra of NC1(84−116) bound to C_4 (solid line), C_8 (dotted line), and C_{18} (dashed line) coated plates.

amplitude of the maxima occur. This is consistent with the suggestion of Jasanoff and Fersht (*31*) that, for a peptide with helical propensity, helicity is generally at a maximum by 20−30% TFE and folding complete by 50%. The presence of an isodichroic point at 203 nm indicates that the peptide undergoes a simple random coil to α-helix transition and that, according to a two-state model, an equilibrium exists between the two conformers. In 41% TFE, the molar ellipticity at 222 nm, after the cooperative coil-to-helix transition, is −23 560 deg cm^2 dmol^{-1}. The corresponding helical content, calculated by the method of Chen et al. (*37*), is 65%.

CD spectra of Figure 2B show that, in the presence of increasing amounts of SDS, the peptide folds into an α-helical conformation. At 0.1 mM SDS (i.e., below the cmc), the CD spectrum shows an increase of the amplitude of the 220−230 broad negative peak. Between 0.2 and 2 mM SDS, the peptide−SDS solutions become turbid and therefore were not analyzed. This aggregation, occurring at low SDS concentration, is likely to be a consequence of the neutralization of positive peptide charges by sulfate groups of SDS, leading to an enhancement of the intermolecular hydrophobic interactions. Formation of SDS−peptide micelles at higher SDS concentrations prevents such interactions

and thus prevents aggregation. For 3 mM SDS (i.e., around the cmc), the maximum helicity is reached, and increasing the amount of SDS up to 100 mM does not change the spectral shape. The slight decrease in amplitude of the minima, observed above 3 mM SDS, may be due to some micelle-dependent light scattering. A helical content of 53% can be calculated from the maximal molar ellipticity of $-19\,250$ deg cm^2 dmol^{-1} at 222 nm. On the whole, the SDS titration of NC1(84−116) indicates a transition from random coil to α-helix conformation.

Figure 2C shows the CD spectra of NC1(84−116) following incubation with quartz plates coated with C_4, C_8, or C_{18} hydrophobic chains. Although the concentration of the peptide bound on the plates cannot be properly quantified, the spectra clearly indicate that NC1(84−116) is able to firmly bind to these hydrophobic surfaces, and the shape of these spectra indicate that the peptide adopts an α-helical conformation upon binding. Assuming that such hydrophobic surfaces mimic the hydrophobic interior of proteins, these data suggest that NC1(84−116) exhibits a clear tendency to adopt an α-helical fold in a protein context.

NMR Spectroscopy. The 2D homo and heteronuclear NMR experiment recorded with NC1(84−116) dissolved in 10 mM sodium phosphate buffer, pH 6.0, containing 50% TFE, yields well-resolved spectra as illustrated by the extracts of TOCSY and NOESY (Figure 3, panels A and B, respectively). The spectra were assigned using the classical sequential assignment method (*40*), the spin systems were identified with DQF−COSY and TOCSY spectra and with the help of ^{1}H-^{13}C HSQC and HSQC-TOCSY spectra in natural abundance. The sequential assigment was performed with the help of the NOESY spectrum obtained at a mixing time of 250 ms. Despite the poor dispersion of the NH and Hα resonances (all NH resonances are in a range of 0.8 ppm, and 23 of 34 Hα resonances are in a range of 0.3 ppm), the sequential attribution of all spin systems was completed (see Figure 3). The proline 90 residue was only observed in a trans-conformation as deduced from the strong cross-peak between Hα of A89 and Hδ of P90. No sign of any other stable NC1(83-l16) conformation was observed in the NMR spectra. Moreover, no conformational exchange was detected in the millisecond time range as checked by ROESY. The identification of all side chain protons of the basic residues of the peptide was difficult due to their large number (6 arginines and 5 lysines) located adjacent to each other along the sequence (R96−R97−K98, K101−R102, and K110−R111). However, ^{1}H-^{13}C heteronuclear experiments allowed the identification of all proton side chain resonances of the basic residues as well as of most of the ^{13}C resonances. The tables of ^{1}H and ^{13}C chemical shifts are available as Supporting Information. The intensities of NOE cross-peaks were defined from the NOESY spectrum recorded at 293 K with a mixing time of 50 ms. The presence of strong sequential amide cross-peaks (Figure 3B) and the fact that sequential dαN(i, $i + 1$) connectivities are much stronger than the corresponding intraresidue connectivities in the fingerprint region (NH-Hα) of the same spectrum (data not shown) are typical of α-helical folding. All of these data indicate that, in 50% TFE, the NC1(84−116) unfolded/folded equilibrium is largely shifted toward the folded state. Thus, the exchange averaging process of conformational interconversion can be neglected, and NOE intensities can be validly correlated with the interproton distances in the folded state.

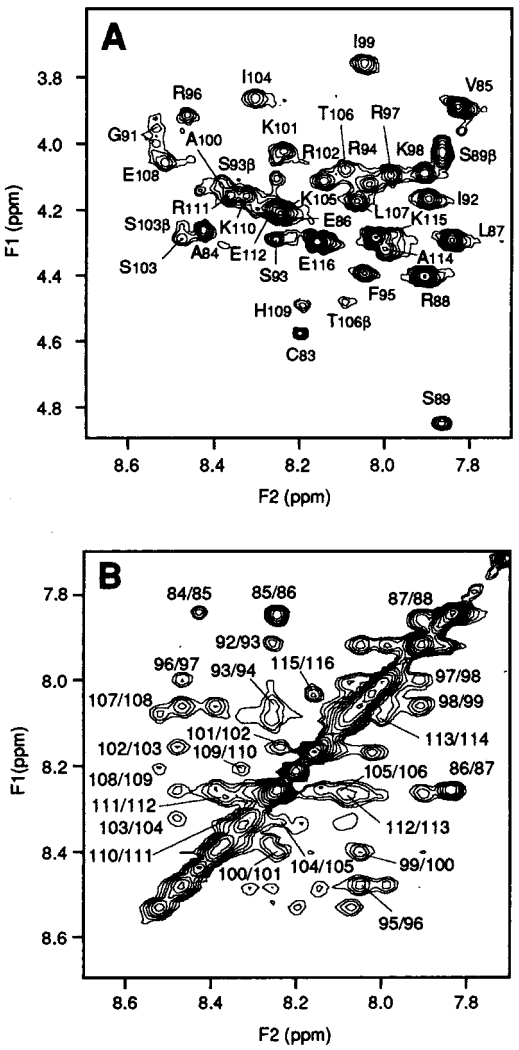

FIGURE 3: Extracts of NMR spectra of NC1(84−116) recorded in 10 mM sodium phosphate buffer pH 6.0 and 50% TFE (v/v) at 293 K. (A) NH-αH region of the TOCSY spectrum recorded with a spin lock of 50 ms. The cross-peaks are labeled by type and residue number. Note that the NH-Hα cross-peak of N113 at 8.08/4.69 ppm was not visible in these conditions. (B) NH-NH region of the NOESY spectrum recorded with a mixing time of 250 ms. The sequential dNN(i, $i + 1$) cross-peaks are labeled by i and $i + 1$ residue numbers.

Secondary Structure. Figure 4 shows the sequence of NC1-(84−116) with an overview of sequential and medium-range NOE connectivities and the chemical shift analysis for ^{1}Hα and ^{13}Cα. Despite the lack of data due to numerous cross-peaks overlapping (indicated by asterisks on the Figure 4), the analysis of NOEs allows the distinction of several structural regions. In the N-terminal part of the peptide (A84-G91 segment), despite the presence of three medium-range connectivities, the strong dαN(i, $i + 1$) and weak dNN(i, $i + 1$) sequential connectivities are the sign of a poorly structured or flexible region. For the I92−S103 segment, strong dNN(i, $i + 1$) and medium dαN(i, $i + 1$) sequential connectivities, and numerous medium dαN(i, $i + 3$), strong dαβ(i, $i + 3$), and weak dαN(i, $i + 4$) medium-range connectivities, are characteristic of an α-helical conformation. For the I104-A114 segment, the presence of dαN(i, $i + 3$), dαβ(i, $i + 3$), and dαN(i, $i + 4$) medium-range connectivities

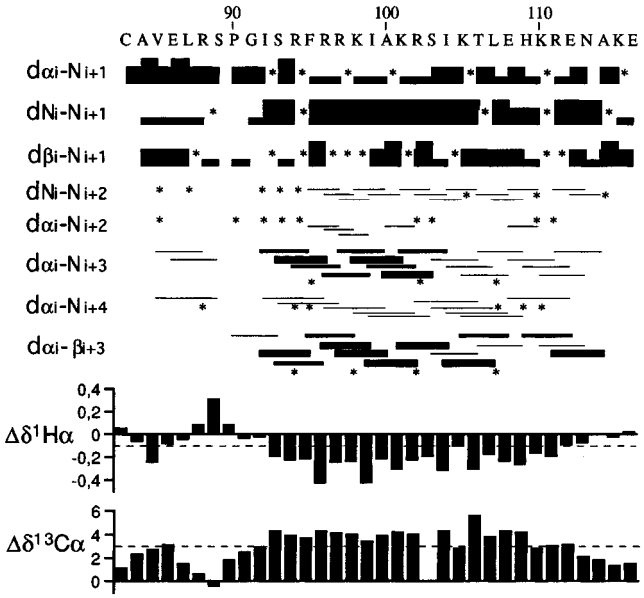

FIGURE 4: Summary of sequential (i, $i + 1$) and medium-range (i, $i + 2$ to i, $i + 4$) NOEs and ^{1}Hα and ^{13}Cα chemical shift differences versus residue number and type. The intensity of NOEs is indicated by the thickness of the connecting bars. Asterisks indicate that the presence of NOEs is not confirmed because of resonances overlapping. ^{1}Hα and ^{13}Cα chemical shift differences were calculated by subtraction of the experimental values to the random coil conformation values reported by Merutka et al. (*46*) and Thanabal et al. (*47*), respectively. The dotted lines indicate the standard threshold values for the α-helix.

Table 1. Statistics on the Final Set of Simulated Annealing Structures of NC1(83−116) in 50% TFE[a]

A. Constraints Used	
distance restraints	
intraresidue	217
sequential	98
medium range	111
total distance restraints	426
B. Statistics for the Final X-PLOR Structures	
number of structures in the final set	19
X-PLOR energy (kcal.mol^{-1})	-169.1 ± 14.2
NOE violations	
number >0.5 Å	none
rms deviation (Å)	0.030 ± 0.007
deviation from idealized covalent geometry	
angles (deg)	1.11 ± 0.07
impropers (deg)	0.22 ± 0.02
bonds (Å)	0.0050 ± 0.0003
rms deviation (Å)	
backbone (C′, C$^\alpha$, N)	
helix (92−108)	0.89
all residues	3.48
all heavy atoms	
helix (92−108)	2.09
all residues	4.79
Ramachandran data[b]	
residues in most favored regions (%)	74.7 ± 5.7
residues in allowed regions (%)	18.8 ± 4.9
residues in generously allowed regions (%)	4.0 ± 3.4
residues in disallowed regions (%)	2.4 ± 1.8

[a] See Experimental Procedures for details of calculation. [b] From PROCHECK (*43*).

clearly indicate the existence of a helical conformation, but the weakness of dαN(i, $i + 3$) connectivities probably indicates some flexibility, or perhaps 3_{10} helix, in this region.

The ^{1}Hα and ^{13}Cα chemical shifts are known to be a straightforward way of analyzing the protein secondary structure, and this has been well-established with the Chemical Shift Index (*44*, *45*). In the case of ^{1}Hα, chemical shifts are not affected by addition of up to 50% TFE (*46*). Relative to a random conformation, an increase of helicity results in an upfield shift of ^{1}Hα resonances, resulting in a negative variation of the chemical shifts ($\Delta\delta^1$Hα). In the central part of the peptide, (S93−R111 segment, Figure 4), the negative $\Delta\delta^1$Hα is equal to or larger than 0.1 ppm, that is, in agreement with the criterion used to define an α-helix conformation (*44*). In the case of ^{13}Cα, chemical shifts are shifted downfield by the formation of an α-helix. However, the threshold of 0.7 ppm used as criterion to assess the presence of α-helix in water is increased to 3 ppm in the presence of TFE. This marked overshift in TFE is mostly due to the influence of this cosolvent on the internal reference TSP (*47*). In NC1(84−116), this 3 ppm threshold for $\Delta\delta^{13}$Cα is reached or exceeded for the I92-E112 segment and confirms the α-helix conformation of this part of the peptide.

Structure Calculation and Analysis. In a first round, only unambiguous NOE constraints were used for structure calculations. An improvement of the constraints set was achieved through NOE back-calculation that allowed the validation of all NOE-derived constraints. Finally, a total set of 426 interproton distance constraints including 98 sequential and 111 medium-range was used for structure calculation (Table 1). Figure 5 A shows the distribution of interproton distance constraints for each residue. It should be noticed that no H-bond constraints were added in the molecular modeling calculation despite the convincing overview of the helical NOE pattern shown in Figure 4. From the set of 50 structures calculated with X-PLOR program (*41*), those having at least one violation of NOE greater than 0.5 Å were discarded as they do not fully satisfy the NMR data. In the final set of the 19 structures selected, pairwise comparison of structures revealed only one structural family. This is illustrated in Figure 6A which shows a view of the 19 structures superimposed from the I92 to the E108 residues.

Figure 6A also shows that an α-helix is well-defined between residues S93 and L107, as also reflected by the local rmsd lower than 0.2 Å and the global rmsd lower than 1 Å. (Figure 5, panels C and B, respectively) Accordingly, the pairwise rmsd for the backbone heavy atoms (C′, C$^\alpha$, N) of the helical region (I92-E108 segment) yields values as low as 0.89 Å (Table 1). The pairwise rmsd for all of the heavy atoms gives a value of 2.09 Å, indicating that the side chains are less well-defined. Both the C and N-terminal parts of the peptide appear to be rather disordered (see Figures 5 and 6). However, it should be noted that the H109-E112 segment was observed as an α-helix in each structure, but failed to superimpose in all structures. This could be related to the weak intensities of medium-range α-helix NOEs (Figure 4) which likely reflect some instability of the helical folding in this region. In addition, in several generated structures, some helical folding was observed on the N-terminal part of the peptide (V85-S89 segment). This is in agreement with the presence of some medium-range NOE constraints in this region (see Figure 4). The stereochemical properties of the backbone dihedral angles provided by the Ramachandran plots showed that more than 93% of the residues are located in the allowed regions (Table 1). The overall energy of the 19 structures is largely negative with values of less than

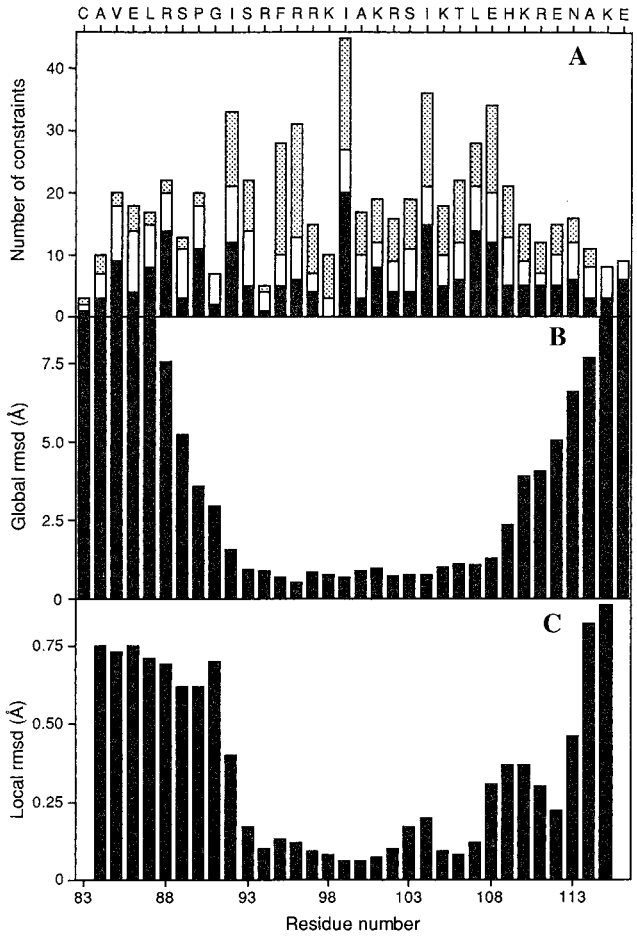

FIGURE 5: Structural characterization of NC1(84−116): number of NOEs and backbone rmsd versus residue number and type. (A) Histogram showing the number of intraresidual constraints (black boxes), sequential i, $i + 1$ constraints (white boxes), and medium i, $i + 2$ to i, $i + 4$ constraints (dotted boxes). Note that each interresidue NOE appears twice, once for each of the two interaction residues. (B) Histogram of the atomic rmsd values for the backbone heavy atoms (N, Cα, and C′) for each residue of the final set of 19 structures. (C) Histogram of the three residue averaged rmsd values for the backbone heavy atoms of the final set of 19 structures.

−169 kcal mol^{-1} (−706 kJ mol^{-1}). The parameters analyzing the deviation from the idealized covalent geometry are homogeneous and fit well with the finding of one family of structures. Hence, an average structure representing this family of structures was calculated and is presented in Figure 6B,C. This Figure shows the distribution of hydrophobic and hydrophilic side chains and reveals that most of basic residues are located on the same side of the helix (Figure 6A,B) and form a twisted basic groove. Conversely, hydrophobic residues are located on the opposite face of the helix.

DISCUSSION

To further characterize and understand the interactions of the NC1 domain of collagen XIV with GAG, we chemically synthesized the NC1(84−116) peptide containing the basic region previously shown to contain a heparin-binding site and to exhibit a conformational change toward helix formation upon heparin binding (*11*). The basic residues critical for heparin binding were identified by analyzing the heparin-binding capacity of purified NC1(84−116) proteolytic cleavage products obtained by treating NC1(84−116) with various proteases. The NC1(84−116) secondary structure in various stabilizing conditions was studied by CD, and its analysis by NMR in 50% TFE allowed us to model its 3D structure.

The analysis of the proteolysis products of NC1(84−116) able to be retained on heparin-agarose media (Figure 1) clearly shows that residues of the segment (97−101), RKIAK, are absolutely required for heparin-binding activity. Interestingly, these residues represent the central part of a XBBBXXBX consensus motif (here FRRKIAKR) which has been proposed as a general heparin-binding consensus sequence (*18*). On that basis, the RRKIAK segment (96−101) probably serves as a primary nucleation site for the recognition of heparin. In addition, the data of Figure 1 indicate that the minimal heparin-binding site of NC1 may extend toward the C-terminus and likely includes at least R102 and K105.

As often observed for peptides, and even domains, when extracted from their native context, the CD analysis reveals that NC1(84−116) is mainly unstructured in aqueous solution. In contrast, it clearly adopts a helical conformation either upon interaction with hydrophobic C4-, C8-, or C18-coated quartz plates or in the presence of TFE or SDS micelles. The latter cosolvents are very useful tools for probing the propensity of peptides for secondary structure formation. However, the major concern over using TFE or SDS is the relevance of the induced peptide structure to the native conformation. Although earlier model studies (*48*) have predicted that all peptide sequences will ultimately become helical in alcohols such as TFE, there is a substantial body of evidence to indicate that this is not always the case. For example, a number of peptides that correspond to β-sheet, loop, and turn regions of proteins are actually disordered in TFE (*28*, *29*, *49*−*51*). In addition, there have been observations of stable β-sheet (*52*−*55*), β-hairpin (*56*−*58*), and even molten globule-like folding intermediates (*59*, *60*) in TFE. It is now recognized that TFE stabilizes α-helical structure only in peptides or protein regions which have an inherent helical propensity (*28*, *49*, *61*, *62*). Further, close examination of TFE-induced non-native structures indicates that they are usually observed only in the presence of high amounts of TFE, are generally highly flexible, and lack stable helical H-bonds (*63*−*65*). By contrast, the use of moderate amounts of TFE allows the exhibition of secondary structure propensities of protein fragments, not only for α-helical but also for β-structure (*65*, *66*).

The precise action of TFE on polypeptide is becoming better understood (*31*, *33*, *60*, *67*, *68*, *70*, and references therein). According to Hirota et al. (*33*), the hydrophobic groups of TFE can associate with each other, and with the hydrophobic groups of peptide residues, in a micelle-like structure. The stabilization of hydrophobic clusters leads to the stabilization of intramolecular hydrogen bonds, to the detriment of the intermolecular H-bonds with water. It can be assumed that, at low concentrations of TFE, such micelle-like structure might mimic the interior of the protein and, thus, would favor the formation of a nativelike peptide conformation. Conversely, increasing the hydrophobicity with high concentrations of TFE would lead to a reduction of the hydrophobic intramolecular interactions between adjacent peptide segments (such as strands in a β-sheet), and thus the corresponding H-bond network would become destabilized. In the case of hydrophobic or amphipathic peptides, this would enhance the formation of α-helix to the

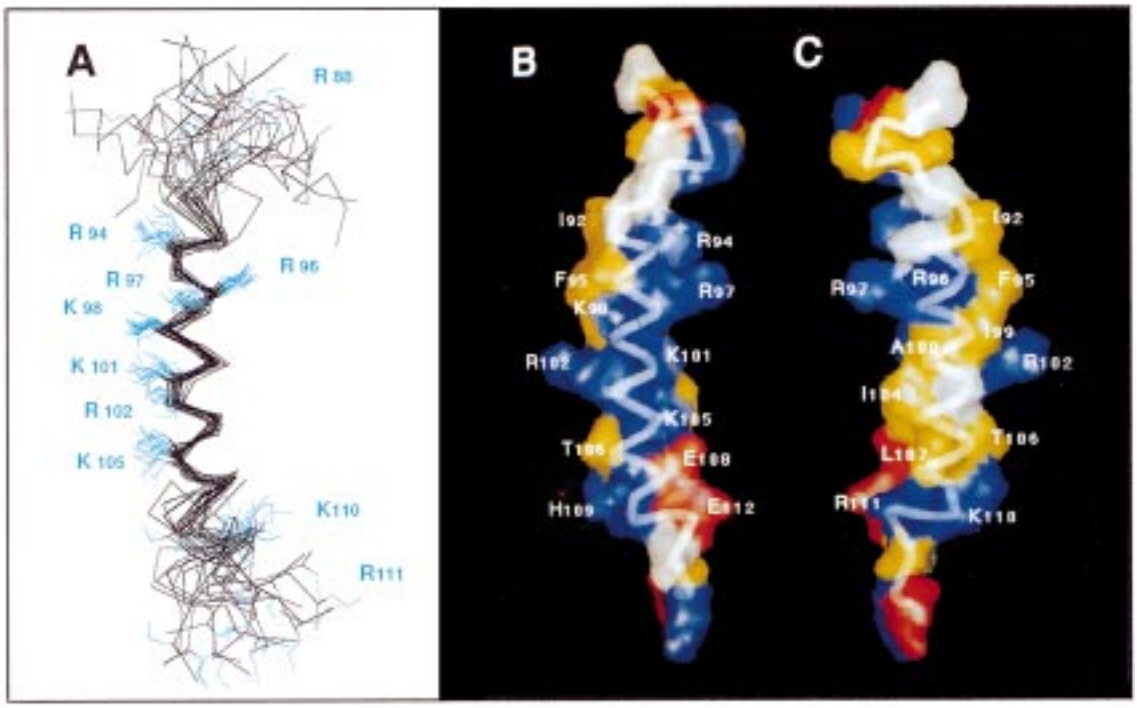

FIGURE 6: Structure of NC1(84−116) calculated from NMR-derived constraints measured in 50% TFE. (A) Superimposition of the α-carbon chains of the 19 final structures aligned for best overlap of residues I92 to E108. The Cα−Cβ and Cβ−Cγ bonds of basic residues are colored in blue. The figure was drawn with Rasmol 2.6 (R. Sayle program). (B and C) Two views of the NMR average structure of NC1(84−116) showing the Connolly's surface distribution of basic (blue), acidic (red), hydrophobic (yellow), and polar (white) residues. The backbone is displayed as a ribbon. The figures were generated with WebLab Viewer 2.0 from MSI.

detriment of β-sheet formation each time a competition between these secondary structures is possible. Consequently, the probing of the conformation tendencies for a peptide should be carefully done by TFE titration (*31*, *66*). On this basis, the fact that the helicity of NC1(84−116) increases continuously over the range 0−40% TFE, and that a clear isodichroic point is observed (indicating that this peptide undergoes a simple two-state random coil/helix transition) leaves no doubt that TFE is only stabilizing the innate α-helical folding of NC1(84−116).

In addition to TFE, SDS provides a hydrophobic environment that mimics either biological membranes or the interior of protein. The mechanism by which micellar and/or nonmicellar SDS acts to induce/stabilize secondary structure in peptides remains poorly understood (*25*, *30*, *32*). However, micellar SDS is well-known to stabilize an α-helical conformation in peptides derived from helical regions of the original protein (*65*, *69*, *71*, and references therein). In the case of NC1(83−116), the CD spectra recorded in micellar SDS and at an SDS concentration around the critical micellar concentration (Figure 2B) revealed a helical folding. This indicates that, in a hydrophobic environment mimicking the protein context, this peptide exhibits a clear propensity to adopt a helical structure.

Recently, a CD procedure was developed to study the conformational changes induced upon interaction of a peptide with hydrophobic C18 chains coated on quartz plates (*34*, *35*). This procedure was later modified to permit the coating of the plates with C4 and C8 chains (*38*). It was shown that, upon interaction with the hydrophobic chains, amphipathic peptides that were unstructured in water adopted either a α-helical or a β-sheet conformation in a sequence-specific and predictable manner. Although this method does not allow a precise quantitation of the extent of the induced conformation, it does constitute a powerful approach toward evaluating the conformational propensities of a peptide by mimicking the hydrophobic interior of a protein. The CD spectra of NC1(84−116) bound to C4-, C8-, or C18-coated plates (Figure 2C) are typical of α-helical folding. Inspection of the spectra reveals that the spectra are all similar, indicating that the peptide adopts essentially the same conformation when bound to each of the surfaces. Clearly the peptide is able to form an amphipathic α-helix stabilized by the interaction between its hydrophobic side chain residues and any hydrophobic surface.

In summary, although NC1(84−116) appears predominantly unstructured in water, all of the conformational changes observed by CD in the different cosolvents (TFE, SDS), or when the peptide interacts with hydrophobic surfaces or with heparin, clearly indicate that the peptide possesses a great propensity toward the formation of an α-helical structure. In addition, all secondary structure prediction methods tested predict the presence of α-helix in this region of NC1 (see Supporting Information). Taken together, these results strongly suggest that, in the biological context, this NC1 region is likely able to adopt an α-helical conformation.

The NMR structure analysis of NC1(84−116) in 50% TFE allowed us to accurately determine the peptide's helical conformation (Figure 6). Examination of the arrangement of the amino acid side chains along the helix reveals that the basic residues are essentially localized on one side of the helix and form a basic groove, while the other residues form a rather hydrophobic groove on the opposite face of the helix. It should be noted that unfavorable electrostatic interactions may occur between arginyl and lysyl side chains

brought into close proximity in the helix. This may contribute to helix destabilization and probably explains why NC1(84−116) appears unfolded in water. The basic residues R97, K98, and K101 (shown to be likely the primary nucleation site for the recognition of heparin), together with the R102 and K105 residues (which also appeared to be important for heparin binding), are close together and constitute the central part of the basic groove. This groove is extended toward the N-terminus by R94 and R96, although the latter is not located on the basic side of the helix. In the C-terminal region, H109 together with K110 and R111, which are located on the opposite side of the helix, appears to extend the basic groove. The positioning of these residues results in a twisted conformation to the basic groove (see models in Figure 6). Although these "capping" residues (i.e., R94, R96, K110, and R111) have not been shown to be essential for heparin binding, they may also contribute to the charge density and it is tempting to speculate that the full length of the twisted basic groove (i.e., the 18 residues from R94 to R111) may actually constitute the NC1 GAG-binding site. Undoubtedly, this site presents features in common with the spatial α-helix model of Margalit et al. (*19*). This model involves an amphipathic helix with most of the positively charged residues on the hydrophilic side in interaction with heparin. In addition, Margalit's model predicts two basic amino acids located at about 20 Å apart, facing opposite directions in the α helix. However, examination of our experimental model shows that the α-helices between R94 and R111 comprise a distance of about 27 Å. This distance would accommodate the binding of a hexasaccharide, and it is not unreasonable to speculate that this may be related to the recognition of a specific GAG motif.

In agreement with the proposal of Margalit et al. (*19*), the presence of this twisted basic groove suggests that GAG wrap around the α-helix through electrostatic interactions between their sulfate and carboxylate groups and the basic peptide groups, forming a coiled coil-like structure. Stabilization of the binding likely involves other attractive interactions such as van der Waals forces and hydrogen bonding. Further, the presence of a hydrophobic groove, highlighted by the NC1-(84−116) model, suggests that the formation of a hydrophobic core is also required to bury the hydrophobic residues for a complete stabilization of the helix. Such a hydrophobic core may be formed by the contribution of additional sequences from other regions of NC1 and/or upon the trimerization of NC1 in collagen XIV. In fact, the NC1 domain immediately follows the triple helical COL1 domain in whole collagen XIV. The collagenous part can contribute to a stable trimeric structure for NC1, whereas the momoner of NC1 expressed in bacteria is monomeric and poorly organized (*11*). An additional consequence of this potential trimerization might be a cooperative action of three NC1 GAG-binding sites to reinforce the overall interaction with GAG. Indeed, cooperativity between GAG-binding sites has already been described for other extracellular matrix proteins such as laminin (*72*, *73*) and fibronectin (*74*) and strongly suggested for collagen V (*24*). However, at present, it is not clear whether the NC1 GAG-binding site is fully folded in "free" collagen XIV or whether the interaction with GAG drives the final folding of this site. Certainly, it has often been reported that carbohydrates can stabilize protein folding by counterbalancing unfavorable clusters of basic charges (*75*). For example, it has been postulated that, in the case of heparin-binding growth factors, the role of heparin binding may be to stabilize the native conformation of the growth factors, thereby facilitating their interaction with cellular receptors (see ref *76*, and references therein). Hence, one of the possible roles of GAG binding may be to stabilize the active conformation of NC1, yielding new interactive properties to collagen XIV or to the isolated NC1 domain.

In conclusion, this detailed structural and functional analysis of NC1(84−116), a peptide containing the NC1 heparin-binding region of collagen XIV, has led to an experimental model of the NC1 heparin-binding site. This is broadly in agreement with the consensus helical heparin-binding model of Margalit et al. (*19*), although the GAG-binding site of NC1 is likely to be longer. The model described in the present study suggests the formation of high-order tertiary or quaternary structures to ensure GAG binding and/or stabilization of the GAG-binding domain. High level expression of a recombinant trimeric minicollagen XIV comprising the COL1 and the NC1 domains of collagen XIV (as previously performed with minicollagen XII in this laboratory, ref *77*) is currently in progress and should help us in further elucidation of the structure−function relationship of the NC1 collagen XIV domain.

ACKNOWLEDGMENT

The authors wish to thank Dr. K. Beck for helpful discussions and C. Combet for help with the structure drawing software.

SUPPORTING INFORMATION AVAILABLE

One figure containing the sequence, the hydrophilic/hydrophobic profiles, and the secondary structure prediction of the NC1 domain of collagen XIV; two tables containing the proton and carbon chemical shifts of the assigned residues of NC1(84−116). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. van der Rest, M., and Garrone, R. (1991) *FASEB J. 5*, 2814−2823.
2. Brown, J. C., and Timpl, R. (1995) *Int. Arch. Allergy Immunol. 107*, 484−490.
3. Prockop, D. J., and Kivirikko, K. I. (1995) *Annu. Rev. Biochem. 64*, 403−434.
4. Gordon, M. K., Gerecke, D. R., Nishimura, I., Ninomiya, Y., and Olsen, B. R. (1989) *Connect. Tissue Res. 20*, 179−186.
5. Shaw, L. M., and Olsen, B. R. (1991) *Trends Biochem. Sci. 16*, 191−194
6. Eyre, D. R., Apone, S., Wu, J. J., Ericson, H., and Walsh, K. A. (1987) *FEBS Lett. 220*, 337−341.
7. van der Rest, M., and Mayne, R. (1988) *J. Biol. Chem. 260*, 555−562.
8. Font, B., Aubert-Foucher, E., Goldschmidt, D., Eichenberger, D., and van der Rest, M. (1993) *J. Biol. Chem. 268*, 25015−25018.
9. Lesage, A., Penin, F., Geourjon, C., Marion, D., and van der Rest, M. (1996) *Biochemistry 35*, 9647−9660.
10. Brown, J. C., Mann, K., Wiedemann, H., and Timpl, R. (1993) *J. Cell Biol. 120*, 557−567.
11. Giry-Lozinguez, C., Aubert-Foucher, E., Penin, F., Deléage, G., Dublet, B., and van der Rest, M. (1998) *Matrix Biol. 17*, 145−149.
12. Ehnis, T., Dieterich, W., Bauer, M., von Lampe, B., and Schuppan, D. (1996) *Exp. Cell Res. 229*, 388−397.

13. Ehnis, T., Dieterich, W., Bauer, M., Kresse, H., and Schuppan, D. (1997) *J. Biol. Chem. 272*, 20414−20419.
14. Ehnis, T., Dieterich, W., Bauer, M., and Schuppan, D. (1998) *Exp. Cell Res. 239*, 477−480.
15. Wälchli, C., Trueb, J., Kessler, B., Winterhalter, K. H., and Trueb, B. (1993*) Eur. J. Biochem. 212*, 483−490.
16. Gerecke, D. R., Foley, J. W., Castagnola, P., Gennari, M., Dublet, B., Cancedda, R., Linsenmayer, T. F., van der Rest, M., Olsen, B. R., and Gordon, M. K. (1993) *J. Biol. Chem. 268*, 12177−12184.
17. Bauer, M., Dieterich, W., Ehnis, T., and Shuppan, D. (1997) *Biochim. Biophys. Acta 1354*, 183−188.
18. Cardin, A. D., and Weintraub, H. J. R. (1989) *Arteriosclerosis 9*, 21−32.
19. Margalit, S., Fischer, N., and Ben-Sasson, S. A. (1993) *J. Biol. Chem. 268*, 19228−19231.
20. Busby, T. F., Argaves, W. S., Brew, S. A., Pechik, I., Gilliland, G. L., and Ingham, K. C. (1995) *J. Biol. Chem. 270*, 18558−18562.
21. Deprez, P. N., and Inestrosa, N. C. (1995) *J. Biol. Chem. 270*, 11043−11046.
22. Mayo, M. H., Ilyini, E., Roongta, V., Dundas, M., Joseph, J., Lai, C. K., Maione, T. E., and Daly, T. J. (1995) *Biochem. J. 312*, 357−365.
23. Wong, P., Hampton, B., Szylobry, E., Gallagher, A. M., Jaye, M., and Burgess, W. H. (1995) *J. Biol. Chem. 270*, 25805−25811.
24. Delacoux, F., Fichard, A., Geourjon, C., Garrone, R., and Ruggiero, F. (1998) *J. Biol. Chem. 273*, 15069−15076.
25. Wu, C.-S. C., Ikeda, K., and Yang, J. T. (1981) *Biochemistry 20*, 566−570.
26. Gierasch, L. M. (1989) *Biochemistry 28*, 923−930.
27. Mammi, S., and Peggion, E. (1990) *Biochemistry 29*, 5265−5269.
28. Segawa, S., Fukuno, T., Fujiwara, K., and Noda, Y. (1991) *Biopolymers 31*, 497−509.
29. Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H. C., Lerner, R. A., and Wright, P. E. (1992) *J. Mol. Biol. 226*, 819−835.
30. Zhong, L., and Johnson, W. C., Jr. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4462−4465.
31. Jasanoff, A., and Fersht, A. R. (1994) *Biochemistry 33*, 2129−2135.
32. Waterhous, D. V., and Johnson, W. C., Jr. (1994) *Biochemistry 33*, 2121−2128.
33. Hirota, N., Mizuko, K., and Goto, Y. (1998) *J. Mol. Biol. 275*, 365−378.
34. Blondelle, S. E., Ostresh, J. M., Houghten, R. A., and Pérez-Payá, E. (1995) *Biophys. J. 68*, 351−359.
35. Steer, D. L., Thompson, P. E., Blondelle, S. E., Houghten, R. A., and Aguilar, M. I. (1998) *J. Pept. Res. 51*, 401−412.
36. Charrier, V., Buckley, E., Parsonage, D., Galinier, A., Darbon, E., Jaquinod, M., Forest, E., Deutscher, J., and Clairbon, A. (1997) *J. Biol. Chem. 272*, 14166−14174.
37. Chen, Y.-H., Yang, J.-T., and Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
38. McLeish, M. J., Hill, J., Wirth, H.-J., and Aguilar, M.-I. (1996) *Protein Sci. 6* (Suppl. 1), 285.
39. Penin, F., Geourjon, C., Montserret, R., Böckmann, A., Lesage, A., Yang, Y. S., Bonod-Bidaud, C., Cortay, J. C., Nègre, D., Cozzone, A. J., and Deléage, G. (1997) *J. Mol. Biol. 270*, 496−510.
40. Wüthrich, K. (1986) in *NMR of proteins and nucleic acids*, John Wiley and Sons, New York.
41. Brünger, A. T. (1992) *Xplor, a system for crystallography and NMR*, Yale University Press, New Haven, CT.
42. Geourjon, C., and Deléage, G. (1995) *Comput. Appl. Biosci. 11*, 681−684.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
44. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
45. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
46. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biol. NMR 5*, 14−24.
47. Thanabal, V., Omecinsky, D. O., Reily, M. D., and Cody, W. L. (1994) *J. Biomol. NMR 4*, 47−59.
48. Thomas, P. D., and Dill, K. A. (1993) *Protein. Sci. 2*, 2050−2065.
49. Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochemistry 31*, 8790−8798.
50. Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., and Wright, P. E. (1993) *Biochemistry 32*, 6337−6347.
51. Shin, H.-C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., and Wright, P. E. (1993) *Biochemistry 32*, 6356−6364.
52. Lu, Z. X., Fok, K. F., Ericson, B. W., and Hugli, T. E. (1984) *J. Biol. Chem. 259*, 7367−7370.
53. Martenson, R. E., Park, J. Y., and Stone, A. L. (1985) *Biochemistry 24*, 7689−7695.
54. Mutter, M., Maser, F., Altmann, K. H., Toniolo, C., and Bonora, G. M. (1985) *Biopolymers 24*, 1057−1074.
55. Narayanan, U., Keiderling, T. A., Bonora, G. M., and Toniolo, C. (1986*) J. Am. Chem. Soc. 108*, 2431−2437.
56. Blanco, F. J., Jimenez, M. A., Pineda, A., Rico, M., Santoro, J., and Nieto, J. L. (1994) *Biochemistry 33*, 6004−6014.
57. Blanco, F. J., and Serrano, L. (1995) *Eur. J. Biochem. 230*, 634−649.
58. Ramirez-Alvarado, M., Blanco, F. L., and Serrano, L. (1996) *Nat. Struct. Biol. 3*, 604−612.
59. Buck, M., Radford, S. E., and Dobson, C. M. (1993) *Biochemistry 32*, 669−678.
60. Alexandrescu, A. T., Ng, Y. L., and Dobson, C. M. (1994) *J. Mol. Biol. 235*, 587−599.
61. Hamada, D., Kuroda, Y., Tanaka, T., and Goto, Y. (1995) *J. Mol. Biol. 254*, 737−746.
62. Shiraki, K., Nishikawa, K., and Goto, Y. (1995) *J. Mol. Biol. 245*, 180−194.
63. Buck, M., Schwalbe, H., and Dobson, C. M. (1995) *Biochemistry 34*, 13219−13232.
64. Schönbrunner, N., Wey, J., Engels, J., Georg, H., and Kiefhaber, T. (1996) *J. Mol. Biol. 260*, 432−445.
65. Najbar, L. V., Craik, D. J., Wade, J. D., Salvatore, D., and McLeish, M. J. (1997) *Biochemistry 36*, 11525−11533.
66. Vieguera, A. R., Jiménez, M. A., Rico, M., and Serrano, L. (1996) *J. Mol. Biol. 255*, 507−521.
67. Rohl, C. A., Chakrabartty, A., and Baldwin, R. L. (1996) *Protein Sci. 5*, 2623−2637.
68. Luo, P., and Baldwin, R. L. (1997) *Biochemistry 36*, 8413−8421.
69. McLeish, M. J., Nielsen, K. J., Wade, J. D., and Craik, D. J. (1993) *FEBS Lett. 315*, 323−328.
70. Cammers-Goodwin, A., Allen, T. J., Oslick, S. L., McClure, K. F., Lee, J. H., and Kemp, D. S. (1996) *J. Am. Chem. Soc. 118*, 3082−3090.
71. McLeish, M. J., Nielsen, K. J., Najbar, L. V., Wade, J. D., Lin, F., Doughty, M. B., and Craik, D. J. (1994) *Biochemistry 33*, 11174−11183.
72. Sung, U., O'Rear, J. J., and Yurchenco, P. D. (1993) *J. Cell Biol. 123*, 1255−1268.
73. Colognato-Pyke, H., O'Rear, J. J., Yamada, Y., Carbonetto, S., Cheng, Y.-S., and Yurchenco, P. D. (1995) *J. Biol. Chem. 270*, 9398−9406.
74. Woods, A., Couchman, J. R., Johansson, S., and Höök, M. (1986) *EMBO J. 5*, 665−670.
75. Wyss, D. F., and Wagner, G. (1996) *Curr. Opin. Biotechnol. 7*, 409−416.
76. Pineda-Lucena, A., Jimenez, M. A., Lozano, R. M., Nieto, J. L., Santoro, J., Rico, M., and Gimenez-Gallego, G. (1996) *J. Mol. Biol. 264*, 162−178.
77. Mazzorana, M., Snellman, A., Kivirikko, K. I., van der Rest, M., and Pihlajaniemi, T. (1996) *J. Biol. Chem. 271*, 29003−29008.

BI9900222