Communication

An Amino Acid Sequence Common to Both Cartilage Proteoglycan and Link Protein*

(Received for publication, May 15, 1985)

Peter J. Neame, Jean-Pierre Perin‡, Francois Bonnet‡, James E. Christner, Pierre Jolliet‡, and John R. Baker

From the Institute of Dental Research and Diabetes Research and Training Center, University of Alabama, Birmingham, Alabama 35294 and the Laboratoire des Protéines, Unité Centre National de la Recherche Scientifique 102 alliée a l’Institut National de la Santé et de la Recherche Médicale, Université de Paris V, 45 Rue des Saints-Pères, F75270 Paris Cedex 06, France

Cartilage proteoglycan monomers associate with hyaluronic acid to form proteoglycan aggregates. Link protein, interacting with both hyaluronic acid and proteoglycan, serves to stabilize the aggregate structure. In the course of determining the primary structure of link protein, two peptides produced by digestion of rat chondrosarcoma link protein with trypsin or chymotrypsin have been selectively purified by immunoaffinity chromatography on a column of monocolonal anti-link protein antibody (8A4) immobilized to Sepharose 4B. These peptides have been sequenced using the double-coupling dimethylaminobenzene isothiocyanate/phenyl isothiocyanate procedure. A consensus sequence, Cys-X-Ala-Gly-Trp-Leu-X-Asp-Gly-Ser-Val-X-Tyr-Pro-Ile-X-X-Pro, obtained by comparing the affinity-isolated trypptic peptide with the affinity-isolated chymotryptic peptide and an overlapping trypptic peptide, shows homology with a sequence obtained from the NH2-terminal of a CNBr peptide from proteoglycan core protein of bovine nasal cartilage: Ser-Ser-Ala-Gly-Trp-Leu-Ala-Asp-Arg-Ser-Val-Arg-Tyr-Pro-Ile-Ser-. We suggest that the common sequence is structurally important to the function of these proteins and may be involved in the binding of both link protein and proteoglycan to hyaluronic acid.

Both proteoglycan (PG) and link protein (LP) (molecular mass 44–48 kDa), which are found in dissociative extracts of cartilage, bind to hyaluronic acid (HA) with high affinity (1). When reaggregated, these three macromolecules form a stable ternary complex with repeating PG-link protein units on an HA backbone which provides cartilage with many of its properties (2). While both LP and PG share the property of binding to HA and therefore might be expected to have some common characteristics, to date there has been no evidence of any similarity in their structures; in fact, the converse seems to be true. Amino acid analyses indicate that the hyaluronic acid-binding region of proteoglycan (HABR) has relatively high levels of glutamic, threonine, proline, methionine, and hexosamine as compared to LP (3, 4), whereas LP has relatively high levels of phenylalanine, tyrosine, and lysine. In addition, antibodies have been prepared which are specific for either PG-HABR or LP but do not cross-react (5). There is little published sequence data which would make a direct comparison between PG and LP possible: 76 residues from bovine nasal cartilage (BNC) LP (6), 20 residues of rat chondrosarcoma LP (RC-LP) (7) and 158 residues from BNC PG, of which 82 are from the HABR (8, 9). None of the available data indicate that there is any similarity between PG and LP.

During the course of determining the primary structure of LP, we have determined two similar sequences which have homology with a sequence derived from bovine nasal cartilage PG (8, 9). The LP-derived sequences were determined from peptides isolated by immunoaffinity chromatography on immobilized monoclonal antibody 8A4, an antibody originally raised to RC-LP and since found to bind avidly to all types of LP described to date regardless of species or tissue of origin (7).

In this communication, we describe the amino acid sequence of a region which is common to RC-LP and BNC-PG. This is the first description of any structural similarity in these two HA-binding glycoproteins. We have utilized RC-LP as it is readily purified to chromatographic homogeneity. The convention used here for describing subtypes of LP is that LP1 and LP2 are the predominant forms found in cartilagenous tissue. LP1 has the higher molecular weight and the greatest degree of glycosylation (10). LP3 is an apparently fully functional form of LP which is prepared by trypsin or clostripain digestion of the PG aggregate.

MATERIALS AND METHODS

HPLC solvents were purchased from Burdick and Jackson, Muskegon, MI. Reagents used in sequence determination were Sequanal grade purchased from Pierce Chemical Co. Trypsin (EC 3.4.14.4) (t1-trypsin-like) and chymotrypsin (EC 3.4.21.1) were obtained from Worthington. The HPLC equipment was either a Rainin Apple microcomputer controlled system with two pumps and a Gilson Holochrome UV monitor or a Perkin-Elmer series 4 and an LC-85 monitor. Columns used for reversed-phase HPLC of peptides were either a Brownlee Aquapore RP-300 or a Vydac C18 protein column (both columns were 4.6 x 250 mm). Other reagents were Baker analytical grade.

Peptide Preparation and Analysis—Link protein was prepared from rat chondrosarcoma by sequential associative and dissociative separations followed by gel filtration as described by Baker and Caterson (10). The LP-containing fractions were precipitated with 50% ethanol and the protein was further purified on Sephacryl S-200 in 4M guanidine HCl, as described by Tang et al. (11). The LP was at least 95% pure by sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis: the major band had a molecular weight of 46,000. The tryptic fragment, T-2200-2 or RC-LP3, was prepared as described by Hascall and Heinegard (12) and was a generous gift from Dr. J. Stevens. The LP was reduced and carboxymethylated as described by Allen (13).

Both tryptic and chymotryptic digestions were performed at an enzyme/substrate ratio of 1:40, in 0.1 M Tris-HCl, 0.1 M guanidine HCl, pH 7.4, at 37 °C for 16 h. The concentration of LP was 0.25 mg/ml. Peptides derived from up to 5 nmol of LP were adjusted to 4 M. 

© 1985 by The American Society of Biological Chemists, Inc. Published in U.S.A.
by the addition of solid guanidine HCl and a sample (up to 2 ml) adsorbed to a reversed-phase HPLC column which had been equilibrated in 0.1% aqueous trifluoroacetic acid at a flow rate of 1 ml/min. Elution was achieved with an ascending gradient of acetonitrile (0.1% trifluoroacetic acid): 0–10% acetonitrile in 10 min and 10 to 40% in 90 min. The eluant was monitored at 214 nm.

The peptide CN-2 RA/6B from BNC-PG was prepared by CNBr cleavage of PG and isolated as described previously (8, 9). Briefly, cyanogen bromide-cleaved PG monomer was fractionated by gel filtration on Sepharose CL-6B and fractions from an included peak (CN-2) were pooled, reduced, and carboxymethylated. These peptides were further fractionated by gel filtration on Sepharose CL-6B. CN-2 RA/6B was homogenous by sodium dodecyl sulfate-polyacrylamide gel electrophoresis prior to and after deglycosylation with HF. It had an apparent molecular weight of 150,000 prior to deglycosylation; this was reduced to 50,000 after HF treatment.

**Sequence Analysis**—Peptide microsequencing was performed on LP-derived peptides by the double coupling dimethylaminosobenzene isothiocyanate method of Chang (14) as modified by Allen (13). Dimethylaminobenzene thiodyantoin amino acids were identified by reversed-phase HPLC and TLC as described by Chang (14). Sequence analysis of CN-2 RA/6B was performed on a Beckman 890C spinning cup sequencer as described previously (8, 9).

**Immunospecific Chromatography**—A column of immobilized monomodal antibody 8A4 was prepared by coupling protein A-purified 8A4 to cyanogen bromide-activated Sepharose (Pharmacia) according to the manufacturer’s instructions. The ligand concentration was 1.3 mg/g wet gel as determined by the loss of absorption at 280 nm after coupling. Immunospecific chromatography was performed by running the peptide mixtures (1 ml), dissolved in 0.1 M sodium chloride, 10 mM Tris-HCl, pH 7.4, onto the bed of the immobilized antibody column (2×9.9 cm). After adsorption for 10 min at room temperature, the unbound peptides were washed off by addition of further buffer (2 bed volumes) and examined by reversed-phase HPLC (Figs. 1 and 3). The column was washed extensively (10 bed volumes) in buffer and then eluted with 3×2-ml aliquots of 1 M acetic acid. The eluate was dried on a Savant Speedvac, dissolved in 4 M guanidine HCl, applied to the reversed-phase HPLC column, and eluted as described above.

**RESULTS AND DISCUSSION**

A chymotryptic digest of RC-LP3 was applied to an 8A4-Sepharose 4B column. A pool of peptides which failed to bind and bound peptides which were eluted with 1 M acetic acid were analyzed by reversed-phase HPLC (Fig. 1). Only two peaks appear in the analysis of peptides which bound to the column (C-1 and C-2, Fig. 1B). The second peak (C-2) appears to be an artifact of the chromatographic system, as it contains little protein by amino acid analysis and is also seen in the affinity column void. The complete amino acid sequence of C-1 is shown in Fig. 2. As expected, there are many peptides (approximately 50) which do not bind to 8A4 and are therefore resolved from the epitope-containing peptide by affinity chromatography (Fig. 1A).

A tryptic digest of RC-LP2 was similarly fractionated on an 8A4-Sepharose 4B column and the bound and unbound peptides were examined by reversed-phase HPLC. The former group of peptides (Fig. 3, lower panel) contained only two major peptides, termed T-1a and T-1b, and several quantitatively minor peptides. Both T-1a and T-1b have the same sequence (Fig. 2) and identical amino acid analyses (not shown), so the appearance of two peaks may be a result of oxidation (i.e., of a tryptophan residue). Among the approximately 40 tryptic peptides which did not bind to the column of immobilized 8A4, one (labeled T-2, Fig. 3, upper panel) possessed a sequence which overlapped the NH$_2$-terminal of C-1 (Fig. 2). The two LP-derived sequences, T-1 and T-2 + C-1 are largely homologous apart from variability at the COOH-terminal and differences as follows: Asp or Asn (residue 7, arbitrarily numbered from the NH$_2$-terminal of T-1), Ser or Ala (residue 12), Gln or Arg (residue 17), and Thr-Lys or Ser-Arg (residues 21–22). The molar yields of C-1, T-1, and T-2 after HPLC separation (approximately 60–75% of the molar quantity of LP applied) is the same as for the majority of peptides. This indicates that these peptides derive from separate regions of the LP molecule and that they are not different gene products.

As illustrated in Fig. 2, there is also a high degree of homology between the LP-derived sequences containing the binding sites for 8A4 (peptides T-1a and T-1b) and the sequence obtained by overlapping C-1 and T-2) and the 16 amino acids determined to be the NH$_2$-terminal of the proteoglycan-derived peptide, CN-2 RA/6B. The conserved sequence is: Ala-Gly-Trp-Leu-X-Asp-X-Ser-Val-X-Tyr-Pro-Ile. The match with the sequence derived exclusively from C-1 and T-2 is more complete. In this case, the only amino acid difference in 14 amino acids is the Gly-Arg change at residue 14 (Fig. 2).

The PG-derived peptide in its unreduced form is disulfide bonded to at least two other cyanogen bromide peptides (9) and a portion is therefore likely to be located in the HABR of PG. It has previously been demonstrated that the NH$_2$-terminal of cartilage chondroitin sulfate PG is at the same end of the molecule as the HABR (8, 15) and therefore the NH$_2$-terminal of CN-2 RA/6B must be in the HABR. Based on its amino acid and hexosamine analysis (8), CN-2 RA/6B spans the keratan sulfate region of PG and may include a fragment of the chondroitin sulfate region of PG. It is tempting to speculate that the conserved sequences may be functional: perhaps involved in binding to HA or they may be part of complementary binding sites between PG and LP.

As HA is a linear molecule, various alternative models could be proposed for binding sites which cover 10 monomers. In one of these, the binding site is formed by one or more parallel amino acid chains which are approximately parallel to the HA. In the other, the protein forms a more convoluted binding site spanning the keratan sulfate region of PG and may include a fragment of the chondroitin sulfate region of PG. It is tempting to speculate that the conserved sequences may be functional: perhaps involved in binding to HA or they may be part of complementary binding sites between PG and LP.
A Common Sequence in Proteoglycan and Link Protein

Fig. 2. Sequences of the LP peptides which bind to 8A4: comparison with a PG-derived peptide (CN-2 RA/6B). The common amino acids are delineated by blocks; amino acids are numbered arbitrarily from the first residue of T-1. The sequence of CN-2 RA/6B is derived from a cyanogen bromide peptide from bovine nasal cartilage proteoglycan core protein and the LP sequence is derived from peptides obtained from RC-LP; as described in the text. Peptides T-1 and C-1 bind to 8A4-Sepharose while peptide T-2 does not. The complete sequences of the LP-derived peptides are shown. The sequence from CN-2 RA/6B is from the NH2-terminal of this 150-kDa BNC-derived glycopeptide.

![Fig. 3. Reversed-phase HPLC of a tryptic digest of RC-LP2. Legend as for Fig. 1.](image)

may be quite large. We speculate that the sequence described in this communication may be involved in a parallel structured binding site of this type. As HA is not a variable molecule, it is likely that the binding site is evolutionarily conserved. This will be further clarified as the primary and secondary structures of these proteins are determined.

Of the three LP peptides which define a sequence homologous with the NH2-terminal of CN-2 RA/6B, two bind to 8A4-Sepharose and one does not. It is thus possible to deduce a limited sequence requirement for this interaction. The 10 amino acids at the NH2-terminal of T-1 are not required as C-1 binds to 8A4 and commences with Leu-Ala-. The epitope for 8A4 must therefore reside wholly or partially within the COOH-terminal 14 residues of T-1 and/or C-1. This section is shared in large measure with a section of the NH2-terminal of CN-2 RA/6B though residues 21-23 (Fig. 2) may be different and residue 14, glycine, is substituted by arginine.

It might be expected that CN-2 RA/6B would cross-react with 8A4. To test this hypothesis, BNC-PG core protein, BNC-HABR and CN-2 RA/6B were bound to nitrocellulose at a similar molar density to LP. Their binding to nitrocellulose was demonstrated with an anti-keratan sulfate monoclonal antibody, 5D4 (16). Relative to LP, they exhibited very low binding of the antibody, both when this was detected with iodinated 8A4 and when the 8A4 was detected with a labeled goat anti-mouse antibody (data not shown). This is not unexpected, as LP has two binding sites for 8A4 (the sequence in peptide T-1 and the sequence defined by peptides T-2 and C-1) and might therefore be expected to bind the antibody tightly.

We have not investigated further areas of homology between LP and PG, but these should become clear from either the amino acid sequence or the gene structure as these are worked out. It is likely that the major portion of these two molecules are different, as evidenced by differences in their amino acid analyses and their immunological specificity, but there may be a substantial structural similarity between them in the regions which bind to HA or to each other.

REFERENCES

1. Hardingham, T. E. (1979) Biochem. J. 177, 237-247
2. Comper, W. D., and Laurent, T. C. (1978) Physiol. Rev. 58, 255-215
3. Tengblad, A. (1979) Biochim. Biophys. Acta 76, 381-389
4. Heinegard, D., and Hascall, V. C. (1974) J. Biol. Chem. 249, 4250-4256
5. Caterson, B., Baker, J. R., Levitt, D., and Paslavy, J. W. (1979) J. Biol. Chem. 254, 9396-9397
6. Férin, J.-P., Bonnet, F., Piron, V., Jollès, J., and Jollès P. (1980) FEBS Lett. 119, 333-336
7. Baker, J. R., Caterson, B., and Christner, J. E. (1983) Limb Development and Regeneration, Part B (Kelly, R. O., Goetinck, P. F., and MacCabe, J. A., eds) p. 17-24, Alan R. Liss, Inc., New York.
8. Bonnet, F., Le Glèdée, S., Férin, J.-P., Jollès, J., and Jollès P. (1983) Biochim. Biophys. Acta 743, 82-90
9. Férin, J.-P., Bonnet, F., Jollès, J., and Jollès (1984) FEBS Lett. 176, 37-42
10. Baker, J. R., and Caterson, B. (1979) J. Biol. Chem. 254, 2387-2393
11. Tang, L. H., Rosenberg, L., Reiner, A., and Poole, A. R. (1979) J. Biol. Chem. 254, 16523-16531
12. Hascall, V. C., and Heinegard, D. (1974) J. Biol. Chem. 249, 4232-4241
13. Allen, G. (1981) Sequeing of Proteins and Peptides (Work, T. S., and Burdon, R. H., eds) Elsevier/North-Holland, New York
14. Chang, J.-Y. (1983) Methods Enzymol. 91, 455-466
15. Stevens, J. W., Oke, Y., Handle, C., Hascall, V. C., Hampton, A., and Caterson, B. (1984) J. Cell. Biochem. 26, 247-259
16. Caterson, B., Christner, J. E., and Baker, J. R. (1983) J. Biol. Chem. 258, 8848-8854

| T-1 | Gly Gly Leu Asp Trp Cys Asn Ala Gly Trp Leu Ser Asp Gly Ser Val Gly Tyr Pro Ile Thr Lys Pro Arg |
| T-2 | Cys Asp Ala Gly Trp Leu Ala Asp Gly Ser Val Arg |
| C-1 | Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile Ser Arg Pro Trp |
| CN-2 RA/6B: | Ser Ser Ala Gly Trp Leu Ala Asp Arg Ser Val Arg Tyr Pro Ile Ser - - - |

Table: Sequences of the LP peptides which bind to 8A4: comparison with a PG-derived peptide (CN-2 RA/6B)

- **T-1**: Gly Gly Leu Asp Trp Cys Asn Ala Gly Trp Leu Ser Asp Gly Ser Val Gly Tyr Pro Ile Thr Lys Pro Arg
- **T-2**: Cys Asp Ala Gly Trp Leu Ala Asp Gly Ser Val Arg
- **C-1**: Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile Ser Arg Pro Trp
- **CN-2 RA/6B**: Ser Ser Ala Gly Trp Leu Ala Asp Arg Ser Val Arg Tyr Pro Ile Ser