Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3766–3770, June 1986
Biochemistry

Complete amino acid sequence of chicken cartilage link protein deduced from cDNA clones

(extracellular matrix/DNA duplication/sequencing/multiple polyadenylylation)

FERENC DEÁK*†, IBOLYA KISS*†, KENNETH J. SPARKS*, W. SCOTT ARGRAVES‡, GREGORY HAMPIKIAN*, AND PAUL F. GOETINCK*‡§¶

*Department of Animal Genetics, The University of Connecticut, Storrs, CT 06268; and ‡La Jolla Cancer Research Foundation, La Jolla, CA 92037

Communicated by M. J. Osborn, March 3, 1986

ABSTRACT **cDNA clones coding for chicken cartilage link protein were isolated and sequenced. The DNA sequence for the entire core polypeptide of the mature link protein and the predicted signal peptide consists of 1065 nucleotides. The deduced primary translation product (355 amino acids) has a molecular mass of 40.7 kDa; the calculated molecular mass of the mature link protein core polypeptide (340 amino acids) is 39.06 kDa. The DNA sequence contains two tandemly arranged repeat sequences that may code for repeated functional domains of link protein involved in binding to hyaluronic acid. The mRNAs for chicken link protein are 6.0, 5.8, and 3.0 kilobase pairs, and the difference between the sizes of the RNA species lies in the 3′ untranslated region.**

Proteoglycan monomers of bovine (1–3) and chicken (4) hyaline cartilage and the Swarm rat chondrosarcoma (5) can interact with hyaluronic acid to form macromolecular aggregates. This interaction is stabilized by one or more link proteins (6–10). Isolated link protein affects proteoglycan aggregate structure (11) and it can bind to either hyaluronic acid (12) or proteoglycan monomer (13).

Link proteins vary in size (14–18), in part as a result of differences in glycosylation (14, 18). Partial amino acid sequences from Swarm rat chondrosarcoma link protein are homologous to those of bovine link protein (19, 20). Whereas no amino acid sequence has been published for chicken cartilage link protein, it has been reported to be closely related to bovine link protein (21). Here we provide the complete amino acid sequence of chicken link protein deduced from cDNA sequences.

MATERIALS AND METHODS

Isolation of cDNA Clones. The isolation of poly(A)$^+$ RNA from 14-day-old chicken embryo sternal cartilage, synthesis, and cloning of cDNA has been described (22). Two cDNA libraries constructed in either pUC8 or pUC9 vectors (at the *Sal* I–*Eco*RI sites) were screened with a synthetic oligonucleotide probe (GARGCNGARCARGCNAARGT) that was deduced from the amino acid sequence of link protein identical in bovine cartilage (19) and rat chondrosarcoma (20) link proteins (Glu-Ala-Glu-Gln-Ala-Lys-Val). The oligonucleotide was synthesized by the phosphoramidite method using a Microsyn (Systec, Minneapolis, MN) DNA synthesizer and was end-labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase. Filters were prehybridized at 37°C in 0.9 M NaCl/90 mM sodium citrate, pH 7/0.1% NaDodSO$_4$/0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin/0.05% sodium pyrophosphate/10% dextran sulfate. After 6 hr, 2 × 10^6 cpm (20 ng) of end-labeled probe was added per 2 ml of prehybridization mixture per filter. After overnight hybridization the filters were washed four times (20 min per wash) at 37°C and twice (30 min per wash) at 45°C in 0.9 M NaCl/90 mM sodium citrate, pH 7/0.05% sodium pyrophosphate.

DNA Sequencing. The nucleotide sequence was determined by using the method of Sanger *et al.* (23). Inserts of cDNA clones were subcloned in M13 phage vectors (24). The cDNA clones were linearized at one vector–insert border, shortened by BAL-31 nuclease, and treated with T4 polymerase to generate blunt ends. After cleavage by a second enzyme at the opposite vector–insert border, the overlapping set of cloned cDNA fragments was inserted into the replicative form of M13 phage cleaved by the second enzyme and either *Sma* I (mp9) or *Hin*cII (mp8). In some instances linearized and denatured plasmid DNA was sequenced directly.

RNA Analysis. RNA was fractionated on 0.8% agarose/2.2 M formaldehyde gels (25), transferred to nitrocellulose filters (Schleicher & Schuell), and hybridized with radioactively labeled probes (26). The final washing step was performed at 68°C with 0.15 M NaCl/15 mM sodium citrate, pH 7/0.1% NaDodSO$_4$/0.05% sodium pyrophosphate.

Labeling of DNA Probes. Isolated DNA fragments were nick-translated (27) or radioactive DNA strands complementary to link protein mRNA were produced by extension of the universal M13 phage sequencing primer using single-stranded M13 phage DNA containing segments of link protein cDNA.

RESULTS

A cartilage cDNA library was screened with the oligonucleotide probe and three clones gave a positive signal. In hybridization tests with poly(A)$^+$ RNA from chicken sterna and calvaria, one clone was positive with both RNAs, one was negative with both, and one was positive only with cartilage RNA. Nucleotide sequencing revealed that this last clone, designated pLPG2, coded for link protein amino acid sequences (see below).

The cDNA library was rescreened with the complete insert of pLPG2. Nine more clones were found with identical restriction patterns to pLPG2 in overlapping regions. All of the clones were in the pUC8 vector. Clone pLPF4 protrudes more into the 5′ direction and clones pLPF8 and pLPH1 covered the link protein gene farther in the 3′ direction. The combined restriction map and sequencing strategy are presented in Fig. 1.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); kb, kilobase pair(s).

†Present address: Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, H6701 Szeged, P.O. Box 521, Hungary.
§Present address: La Jolla Cancer Research Foundation, 10901 North Torrey Pines Road, La Jolla, CA 92037.
¶To whom correspondence should be addressed.

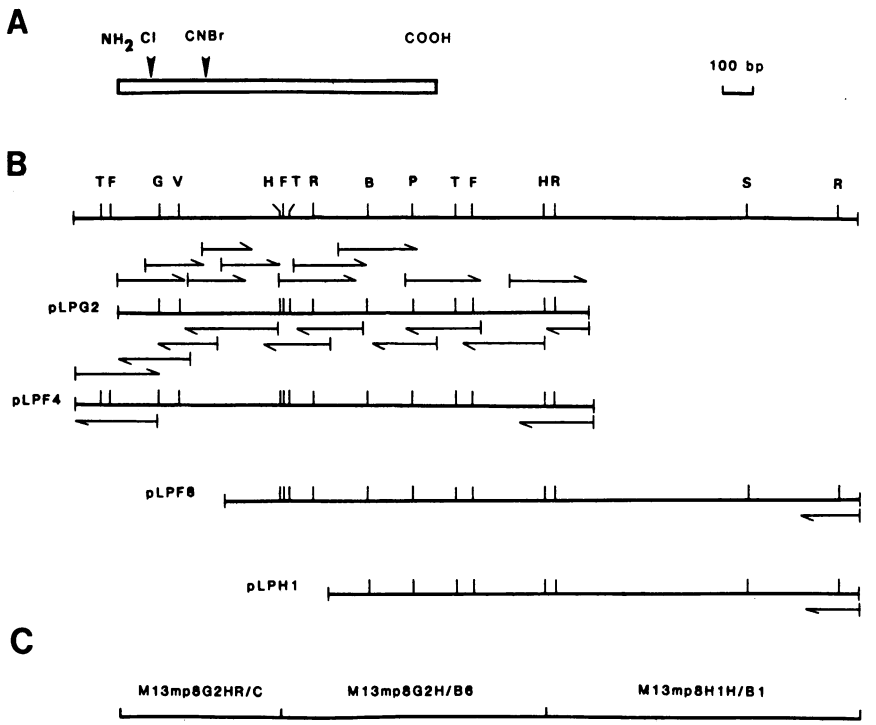

FIG. 1. Composite restriction map and sequencing strategy of cDNA clones coding for chicken cartilage link protein. (*A*) Schematic representation of the core polypeptide of chicken link protein deduced from the nucleotide sequence. Cl and CNBr denote the sites homologous to clostripain and cyanogen bromide cleavage sites of bovine link protein. bp, Base pairs. (*B*) Composite restriction map of the cDNA and sequencing strategy of four recombinants. All cleavage sites are presented for B, *Bgl* I; C, *Sac* I; F, *Hin*fI; G, *Bgl* II; H, *Hin*dIII; P, *Pst* I; R, *Rsa* I; T, *Taq* I; V, *Pvu* II. Only one site is presented for S, *Sau*96I. Arrows indicate the 3′ beginning end of fragments subcloned in M13 phages and the length of the sequence read. (*C*) Subclones used in RNA blot hybridization analysis.

Both strands of the insert of pLPG2 and the 5′ end of the insert of pLPF4 have been sequenced. The combined nucleotide sequence and deduced amino acid sequence are presented in Fig. 2. Within the longest open reading frame the distance between the first codon for methionine and the stop codon is 1065 bp. This sequence codes for a polypeptide of 355 amino acids with a calculated molecular mass of 40,746.7 Da. The deduced polypeptide contains sequences homologous to the partial amino acid sequences known for bovine cartilage (19) and rat chondrosarcoma (20, 28) link proteins.

The first 15 amino acids of chicken link protein core polypeptide are taken to represent a signal peptide. This sequence consists mostly of hydrophobic residues (Fig. 3*A*) and it is homologous to the signal peptide of chicken conalbumin (ref. 30; Fig. 3*B*). Without the signal peptide, the calculated molecular mass of the core polypeptide of link protein is 39,066.6 Da. Except for the signal peptide, the plot of hydropathy (Fig. 3*A*) shows an overall hydrophilic nature (grand average hydropathy score = −0.38; ref. 29). There are five regions where a cluster of positively charged residues occurs (76–84; 101–116; 244–248; 316–325; 341–344) and one region with a negatively charged cluster (133–149) (Fig. 3*C*).

There are three asparagines on the amino-terminal side of the methionine at position 99, which corresponds to the CNBr cleavage site of bovine link protein (18). Two of these satisfy the consensus Asn-Xaa-Ser/Thr N-glycosylation acceptor sequence. None of the asparagines on the carboxyl side of residue 99 fits this consensus sequence, in agreement with the observation that this portion of bovine link protein is not glycosylated (18). The deduced sequence of the chicken link protein contains 11 cysteine residues. One cysteine is in the signal peptide and 10 are in the mature protein.

A repeat sequence analysis of the nucleotide sequence for link protein core polypeptide revealed a high degree of homology between nucleotides 514–762 and 808–1053 (Fig. 4). These two sequences show an overall homology of 58.02%. This is reflected in a 47.62% homology between amino acid sequences for residues 172–254 and 270–351 (Fig. 4). The two tandemly repeated sequences contain 8 of the 10 cysteine residues found in the mature link protein core polypeptide. Within the two tandemly arranged repeats there are highly conserved regions between nucleotides 616–681 and 913–978 with a homology of 68.18%. At the protein level these highly conserved regions are between amino acid residues 206–227 and 305–326 and they also have a homology value of 68.18%. The flanking amino acids of the highly conserved region are cysteine residues.

In addition to the translated region, the cDNA clones cover 136 bp of the 5′ untranslated region and ≈1400 bp of the 3′ untranslated region. Except for slight variations in the number of adenosines, the 3′ ends of pLPG2 and pLPF4 are identical as are the ends of pLPF8 and pLPH1. There is one consensus poly(A) signal sequence (AATAAA) at the end of pLPG2 and pLPF4, and two more were found within the last 200 bp of pLPF8 and pLPH1. pLPF8 has not been sequenced in the region that corresponds to the 3′ end of pLPF4 and it may not contain the adenosine-rich region found at the 3′ end of pLPF4.

When pLPG2 insert was hybridized with poly(A)$^+$ RNA from embryonic sterna, 6.0- and 5.8-kilobase-pair (kb) transcripts and a less abundant 3.0-kb transcript were detected (Fig. 5). The two largest messages can best be resolved when the electrophoresis is prolonged. When three nonoverlapping fragments of pLPG2 and pLPH1 were used as probes, the same three bands were detected. When a genomic clone fragment located downstream from the region hybridizing with pLPH1 was used as a probe, only the large mRNA species were detected (Fig. 5, lane 8). Therefore, the difference between the large and small poly(A)$^+$ RNA species lies in the 3′ untranslated region.

DISCUSSION

cDNA clones for chicken cartilage link protein were found by using a mixed synthetic oligonucleotide probe whose construction was based on bovine and rat link protein sequences. Hybridization occurred even though there were mismatches at positions 2 and 19 of the probe. The nucleotide sequence of clone pLPG2 revealed that it coded for chicken cartilage link protein.

Clones pLPF4, pLPF8, and pLPH1 hybridize with the insert of pLPG2 and the restriction sites in their inserts correspond exactly to those of pLPG2. Therefore, these four inserts are copies of mRNA that codes for the same protein.

FIG. 2. Nucleotide sequence of the cDNA for chicken cartilage link protein. Nucleotide residues are numbered beginning with the first residue of the ATG triplet encoding the initiating methionine. The deduced amino acid sequence is numbered from the initiating methionine. *, Potential asparagine-linked glycosylation sites. Translation stop codons are underlined; AATAAA sequences found in the untranslated regions are boxed. B, bovine cartilage link protein sequence (19), aligned to increase the number of matches. One unidentified bovine amino acid residue (X) was omitted from between amino acid positions 117 and 118. R, rat chondrosarcoma sequence (20, 28). After nucleotide position 1573, the numbering is approximate and the reading frame is arbitrary.

This was confirmed when three subcloned nonoverlapping fragments were shown to hybridize to cartilage mRNAs of the same size.

The methionine in position 1 is taken to be the first amino acid of the primary translation product of link protein mRNA since it is located within the expected range from the clostripain cleavage site (18) and, furthermore, upstream in the cDNA of the same reading frame there are two stop codons. The calculated molecular mass of 40,746.7 Da for the primary translation product agrees well with the M_r 41,000 value of a chondrocyte cell-free translation product bound by an antibody against link protein (31).

A hypothetical fragment of chicken link protein between the methionine at position 99 and the carboxyl terminus is 257 amino acids long (29,480.8 Da). This value agrees well with the M_r 27,500 value for the corresponding fragment of bovine link protein (18). The distance between the presumed clostripain site and methionine residue 99 is 59 amino acids in chicken link protein (6640.5 Da). This methionine corresponds to the only methionine of bovine link protein. Even if one were to add 4000 Da for the estimated amount of carbohydrate in bovine link protein, the value obtained is still much smaller than the M_r 18,000 value published for that bovine link protein fragment (18). The latter value may therefore be an overestimate due to the influence of the carbohydrate on the electrophoretic mobility of link protein.

Eleven cysteines were found in the deduced primary sequence. The first one is contained in the secretory signal peptide that is cleaved off, and the remaining cysteines may all form disulfide bridges, in agreement with the estimated five or six disulfide bridges and an absence of cysteine sulfhydryl groups for bovine link protein (14). CNBr cleavage of bovine link protein does not produce fragment separation until the molecule is reduced (18). This suggests that one or more disulfide bridges span the CNBr site. Since the cysteine at position 61 is the only cysteine on the N-terminal side of the two methionines of chicken link protein, there can only be one disulfide bridge spanning the CNBr cleavage sites.

A repeat sequence search of the link protein coding sequence revealed two tandemly arranged homologous sequences that code for amino acid repeats consisting of 83 and 82 residues each. Within each of these repeats there is a highly conserved region consisting of 22 amino acids. Chicken amino acid sequence 202–224 of the first repeat is identical to a 23 amino acid sequence of the 24 amino acid T-I peptide of rat chondrosarcoma link protein and chicken amino acid

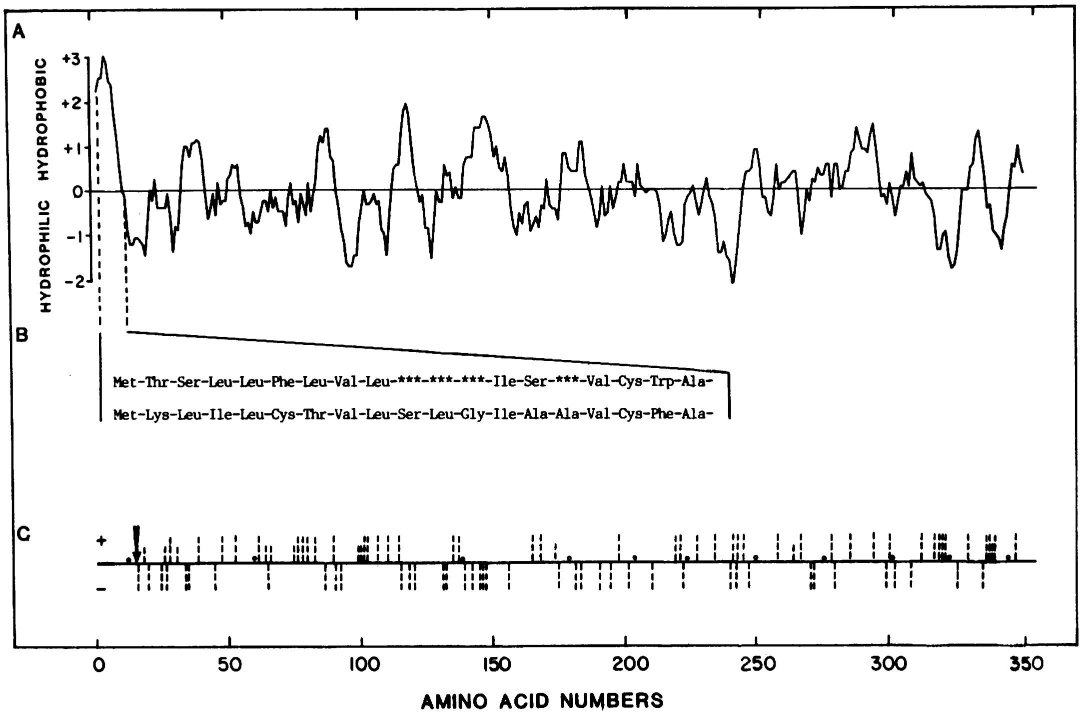

FIG. 3. (*A*) Plot of hydropathy generated according to the method of Kyte and Doolittle (29), using a window of 9 amino acids. The points are plotted over the fifth amino acid. (*B*) Comparison of the sequence of the first 15 amino acids of the deduced primary translation product for chicken link protein core polypeptide with the signal peptide of chicken conalbumin (30). Spaces (***) have been inserted in the link protein sequence to optimize homology. (*C*) Distribution plot of charged residues and cysteine residues of chicken link protein core polypeptide. Upward-projecting, full-length bars denote arginine and lysine residues, half-length bars denote histidines, and downward-projecting bars denote aspartic acids and glutamic acids. Cysteine residues are indicated by open circles along the baseline. The arrow points to amino acid 16, the N-terminal amino acid of the mature link protein core polypeptide.

sequence 305–322 of the second repeat is identical with an 18 amino acid sequence derived from the combined 19 amino acid sequence of the T-2 and C-I peptides of rat chondrosarcoma link protein (28). This observation would suggest that rat chondrosarcoma link protein, like chicken link protein, contains tandemly repeated sequences in this part of the molecule.

Bovine proteoglycan core protein contains an amino acid sequence in its hyaluronic acid binding region that is very similar to link protein sequences (28). The sequence of the last 14 amino acids (Ala . . . Ser) of the 16 amino acid bovine proteoglycan core protein peptide CN-2 RA/6B (28) is more homologous to the sequence of the second repeat of chicken link protein (92.8% homology) than that of the first one (71.4% homology). It was suggested by Neame *et al.* (28) that the sequence similarity between bovine proteoglycan core protein and rat chondrosarcoma link protein may reflect a functional similarity involving binding to hyaluronic acid. If this is the case, then the present results suggest that binding of link protein to hyaluronic acid might be of a bivalent nature involving tandemly repeated functional domains. As a result, two methods of interaction with hyaluronic acid are possible. Either the two domains of link protein interact linearly along a single hyaluronic acid polymer or they bind in an interpolymeric fashion with two hyaluronic acid polymers. The second method of interaction might contribute to the establishment of three-dimensional structures in the matrix by cross-linking of hyaluronic acid polymers, an organization that might be reinforced if proteoglycan core protein would also contain tandemly repeated hyaluronic acid binding

```
                                                                                       (514)  AAC CTA AAC TTC
                                                                                       (808)  --G --G -C- -AT

CAT GAG GCT CAG CAA GCT TGC CTG GAC CAG GAC TCC ATC ATT GCC TCC TTC GAC CAG CTC TAC GAG GCC TGG AGG *** TCA GGG
G-- --A --C GT- --G --C --- --- A-G G-T -G- G-T CAG --- --- AAG G-T -GG --- A-A -T- -CT --- --- -A- CTC CTT --T

CTG GAC TGG TGC AAT GCT GGC TGG CTC AGT GAT GGT TCA GTG CAG TAC CCT ATC ACC AAG CCC AGA GAG CCC TGT GGA GGG AAG
TAT --- C-C --T G-- --C --- --- --G GCA --C --C AGC --C -GC --- --C --- T-- -GA --- --- A-- -G- --C A-C CCC ***

AAT ACG GTG CCC GGT GTC AGA AAC TAT GGC TTC TGG GAT AAA GAG AGG AGC CGA TAT GAT GTT TTC TGC TTT  (762)
--C GA- -CT G-- *** --- C-C TTT GTA --- --T CCT --- --- A-- CAC -AG -TG --- -G- --C -A- --T --C (1053)


                                                                                       (172)  Asn Leu Asn Phe
                                                                                       (270)  Lys Leu Thr Tyr

His Glu Ala Gln Gln Ala Cys Leu Asp Gln Asp Ser Ile Ile Ala Ser Phe Asp Gln Leu Tyr Glu Ala Trp Arg *** Ser Gly
Asp Glu Ala Val Gln Ala Cys Leu Lys Asp Gly Ala Gln Ile Ala Lys Val Gly Gln Ile Phe Ala Ala Trp Lys Leu Leu Gly

Leu Asp Trp Cys Asn Ala Gly Trp Leu Ser Asp Gly Ser Val Gln Tyr Pro Ile Thr Lys Pro Arg Glu Pro Cys Gly Gly Lys
Tyr Asp Arg Cys Asp Ala Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile Ser Arg Pro Arg Lys Arg Cys Ser Pro ***

Asn Thr Val Pro Gly Val Arg Asn Tyr Gly Phe Trp Asp Lys Glu Arg Ser Arg Tyr Asp Val Phe Cys Phe  (254)
Asn Glu Ala Ala *** Val Arg Phe Val Gly Phe Pro Asp Lys Lys His Lys Leu Tyr Gly Val Tyr Cys Phe  (351)
```

FIG. 4. Comparison of the nucleotide sequences 514–762 and 808–1053 and amino acid sequences 172–254 and 270–351 of the tandem repeats of chicken link protein core polypeptide. In the nucleotide sequence, homologies are indicated by "---". Homologies in the amino acid sequences are enclosed by boxes. The regions of very high homology are enclosed with the thicker line. Spaces (***) have been inserted to align for optimal homology.

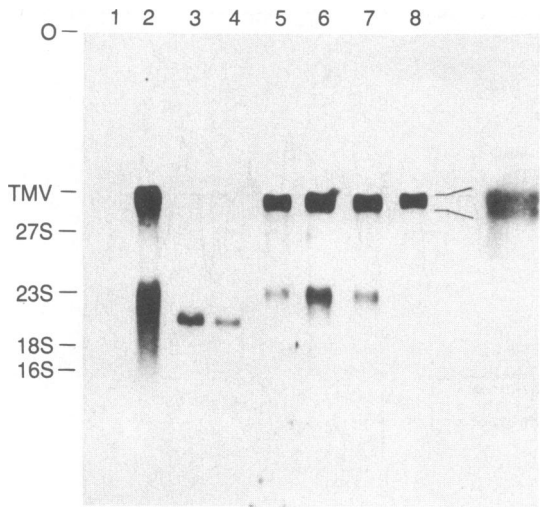

FIG. 5. Blot hybridization analysis of poly(A)$^+$ RNA isolated from chicken embryo calvaria (lanes 1 and 3) or sterna (lanes 2 and 4–8). One microgram of RNA was separated on 0.8% agarose/2.2 M formaldehyde gels, blotted, and hybridized to radioactive probes as follows: lanes 1 and 2, isolated insert of pLPG2; lanes 3 and 4, pA1, a clone coding for chicken actin (34); lane 5, mp8G2HR/C; lane 6, mp8G2H/B6; lane 7, mp8H1H/B1; lane 8, λgLP39.23 H/D, a genomic clone fragment located downstream from the region hybridizing with pLPH1 (unpublished data). For probes used in lanes 5–7, see also Fig. 2. Size markers used were tobacco mosaic virus RNA (TMV, 6.4 kb), chicken 27S (4.6 kb) and 18S (1.8 kb) rRNA, and *Escherichia coli* 23S (2.9 kb) and 16S (1.5 kb) rRNA. O, origin.

domains. An interpolymeric hyaluronic acid interaction of link protein might also help explain the observation that the measured length of hyaluronate in link protein-containing aggregates is five times longer than in the link protein-free aggregates and that the presence of link protein increases the number of monomers per aggregate (11). Modifications of arginine, lysine, and tryptophan residues in the proteoglycan monomer have been shown to abolish aggregate formation (32). In view of the structural similarities of the proposed hyaluronic acid binding regions of link protein and proteoglycan monomer (28), the interaction of link protein with hyaluronic acid might be expected to involve the same amino acids. The two proposed hyaluronic acid binding domains of link protein contain these three amino acids but in different proportions (Fig. 4).

Two major size classes of messages for link protein were observed. The large messages are six times as long as the translated region and the shorter ones are roughly three times as long. The basis for the difference is shown to be at the 3′ untranslated region of the message, most likely resulting from multiple termination and poly(A) events as shown for collagen (33).

We are grateful to Dr. J. Kelley for her help with the analysis of the hydropathy of link protein. This work was supported by Grant HD 09174 from the National Institute of Child Health and Human Development. The paper is Scientific Contribution no. 1143 of the Storrs Agricultural Experiment Station, The University of Connecticut.

1. Hardingham, T. E. & Muir, H. (1972) *Biochim. Biophys. Acta* **279**, 401–405.
2. Hascall, V. C. & Heinegård, D. (1974) *J. Biol. Chem.* **249**, 4232–4241.
3. Rosenberg, L., Hellmann, W. & Kleinschmidt, A. K. (1975) *J. Biol. Chem.* **250**, 1877–1883.
4. Royal, P. D., Sparks, K. J. & Goetinck, P. F. (1980) *J. Biol. Chem.* **255**, 9870–9878.
5. Oegema, T. R., Jr., Hascall, V. C. & Dziewiatkowski, D. D. (1975) *J. Biol. Chem.* **250**, 6151–6159.
6. Hascall, V. C. & Sajdera, S. W. (1969) *J. Biol. Chem.* **244**, 2384–2396.
7. Gregory, J. D. (1973) *Biochem. J.* **133**, 383–386.
8. Heinegård, D. & Hascall, V. C. (1974) *J. Biol. Chem.* **249**, 4250–4256.
9. Hardingham, T. E. (1979) *Biochem. J.* **177**, 237–247.
10. Tang, L.-H., Rosenberg, L., Reiner, A. & Poole, A. R. (1979) *J. Biol. Chem.* **254**, 10523–10531.
11. Buckwalter, J. A., Rosenberg, L. C. & Tang, L.-H. (1984) *J. Biol. Chem.* **259**, 5361–5363.
12. Oegema, T. R., Jr., Brown, M. & Dziewiatkowski, D. D. (1977) *J. Biol. Chem.* **252**, 6470–6477.
13. Caterson, B. & Baker, J. (1978) *Biochem. Biophys. Res. Commun.* **80**, 496–503.
14. Baker, J. R. & Caterson, B. (1979) *J. Biol. Chem.* **254**, 2387–2393.
15. Caputo, C. B., MacCallum, D. K., Kimura, J. H., Schrode, J. & Hascall, V. C. (1980) *Arch. Biochem. Biophys.* **204**, 220–233.
16. Caterson, B. & Baker, J. R. (1979) *J. Biol. Chem.* **254**, 2394–2399.
17. McKeown-Longo, P. J., Velleman, S. G. & Goetinck, P. F. (1983) *J. Biol. Chem.* **258**, 10779–10785.
18. Le Gledic, S., Périn, J.-P., Bonnet, F. & Jollès, P. (1983) *J. Biol. Chem.* **258**, 14759–14761.
19. Périn, J.-P., Bonnet, F., Pizon, V., Jollès, J. & Jollès, P. (1980) *FEBS Lett.* **119**, 333–336.
20. Baker, J. R., Caterson, B. & Christner, J. E. (1983) in *Limb Development and Regeneration*, eds. Kelley, R. O., Goetinck, P. F. & MacCabe, J. A. (Liss, New York), pp. 17–24.
21. Neame, P. J., Christner, J. E. & Baker, J. R. (1984) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **43**, 1695a (abstr.).
22. Deák, F., Argraves, W. S., Kiss, I., Sparks, K. J. & Goetinck, P. F. (1985) *Biochem. J.* **229**, 189–196.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
24. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
25. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16**, 4743–4751.
26. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
27. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
28. Neame, P. N., Périn, J.-P., Bonnet, F., Christner, J. E., Jollès, P. & Baker, J. R. (1985) *J. Biol. Chem.* **260**, 12402–12404.
29. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
30. Thibodeau, S. N., Lee, D. C. & Palmiter, R. D. (1978) *J. Biol. Chem.* **253**, 3771–3774.
31. Treadwell, B. J., Mankin, D. P., Ho, P. K. & Mankin, H. J. (1980) *Biochemistry* **19**, 2269–2273.
32. Hardingham, T. (1981) *Biochem. Soc. Trans.* **9**, 489–497.
33. Meyers, J. C., Dickson, L. A., deWet, W., Bernard, M. P., Chu, M. L., Diliberto, M., Pepe, G., Sangiorgi, F. O. & Ramirez, F. (1983) *J. Biol. Chem.* **258**, 10128–10135.
34. Cleveland, D. W., Lopata, M. A., MacDonald, R. J., Cowan, N. J., Rutter, W. J. & Kirschner, M. W. (1980) *Cell* **20**, 95–105.