The expression of functional link protein in a baculovirus system: analysis of mutants lacking the A, B and B' domains

Judy GROVER and Peter J. ROUGHLEY*

Genetics Unit, Shriners Hospital for Crippled Children, and Department of Surgery, McGill University, Montreal, Canada, H3G 1A6

Functional recombinant human link protein has been produced using a baculovirus expression system. In addition to the intact link protein, three mutant forms have also been expressed. Each mutant bears a deletion equivalent to the protein encoded by one exon in the gene. These deletions represent the A domain, which is thought to be responsible for interaction with aggrecan, and the B or B' domains, which are associated with the interaction with hyaluronate. Such deletions split codons spanning exon boundaries, but maintain the reading frame of the protein and result in the correct amino acid being present at the splice junction. All the recombinant proteins appear as two components

upon SDS/PAGE, though the abundance of the two forms does vary between preparations, as a result of variable substitution by N-linked oligosaccharides. The recombinant intact link protein was able to interact with both hyaluronate and aggrecan, showing that the baculovirus system is able to produce functional molecules. All of the recombinant mutant link proteins were also able to interact with hyaluronate, indicating that both the B and B' domains can function independently. The recombinant mutant link proteins were also able to interact with aggrecan, with the exception of the mutant lacking the A domain, confirming that this ability resides entirely within this domain.

INTRODUCTION

Hyaline cartilages are characterized by the presence of large proteoglycan aggregates that consist of a central hyaluronate filament surrounded by numerous molecules of the proteoglycan aggrecan [1]. In the interterritorial matrix of the cartilage, the hyaluronate filaments form a latticework between neighbouring collagen fibrils [2], which serves to localize the aggrecan in these sites. These proteoglycan aggregates also possess a third component, link protein, which is thought to stabilize the interaction of each aggrecan molecule with hyaluronate [3].

In the human, the link protein gene resides on chromosome 5 [4], and encodes a primary translation product of 354 amino acids [5]. Removal of the leader sequence gives rise to a mature protein of 339 amino acids. The rat link protein is of the same size and bears a high degree of similarity in amino acid sequence, including the conserved positions of 10 cysteine residues [6]. These cysteine residues are all involved in disulphide bonds, and serve to divide the molecule into three domains. Analogous domains are present within the N-terminal globular region of aggrecan [7], the region responsible for the interaction of the proteoglycan with hyaluronate [8]. The link proteins from other species are also well conserved, with the chick showing the greatest variation at present, having 340 amino acids in its mature form [9]. In the chick, the link protein gene has been shown to consist of at least 5 exons, with the translated sequence being contained within 4 exons, whose boundaries flank the disulphide-bonded domains [10].

Most species studied to date possess two consensus sequences for N-linked oligosaccharide substitution in link protein. Linkage occurs at amino acid residues 6 and 41 in the N-terminal region adjacent to the first disulphide-bonded domain [4]. The exception is the rat link protein, which has only the second glycosylation

site [6]. This second site is encoded by the same exon as the first disulphide-bonded domain, whereas the first site is encoded by a separate exon [10]. The site at asparagine residue 41 always appears to be occupied, whereas the site at asparagine residue 6 shows a variable degree of substitution, giving rise to structural heterogeneity in the intact link protein [11]. The intact link protein (LP) with two N-linked oligosaccharides has been termed LP1, and that with one N-linked oligosaccharide has been termed LP2. Smaller forms of link protein accumulate with age as a result of proteolytic processing between the two sites of oligosaccharide substitution, and these have been termed LP3 [12,13].

Proteoglycan aggregates commonly possess link protein and aggrecan in a 1:1 molar ratio [14,15], and the molecules appear to alternate along the hyaluronate filament [16]. This arrangement is favoured by the ability of link protein to interact independently with both the N-terminal globular region of aggrecan and the hyaluronate [17–19]. At neutral pH, the interaction with hyaluronate is reported to be stronger than that with aggrecan [20], whereas under conditions of low pH, the interaction of link protein with aggrecan appears to be more stable than that of either component with hyaluronate [21]. The interaction of link protein with aggrecan and hyaluronate is mediated by its disulphide-bonded domains. The N-terminal of these, termed the A domain, appears responsible for the interaction with aggrecan [22]. The remaining two tandem repeats, termed the B and B' domains, appear responsible for the interaction with hyaluronate [22,23]. However, it has been suggested that the protein sequence of the tandem repeat domains is more compatible with a protein–protein interaction than that of the A domain [24].

The purpose of this work was twofold: first, to determine whether a baculovirus expression system can be used to produce functional recombinant human link protein; and second, to

Abbreviations used: AcNPV, *Autographica californica* nuclear polyhedrosis virus; FBS, fetal bovine serum; HABP, hyaluronate-binding protein; LP, link protein; MMuLV-RT, Moloney murine leukaemia virus reverse transcriptase; pB2, pBlueBac2.

* To whom correspondence should be addressed: Dr. P. J. Roughley, Genetics Unit, Shriners Hospital for Crippled Children, 1529 Cedar Avenue, Montreal, Quebec, Canada, H3G 1A6.

express mutant forms of link protein lacking the A, B or B' domains, to determine how such a deficit affects the ability of the molecule to interact with aggrecan and hyaluronate.

METHODS

Materials

The MAXBAC Baculovirus Expression System was obtained from Invitrogen, through BDH, Montreal, Quebec, Canada. Grace's insect cell medium, Sf-900 serum-free insect cell culture medium, yeastolate, lactalbumin hydrolysate, and Moloney murine leukaemia virus reverse transcriptase (MMuLV-RT) were from GIBCO-BRL, Burlington, Ontario, Canada. Taq DNA polymerase was from Perkin-Elmer Cetus, Montreal, Quebec, Canada. Restriction endonucleases were from New England Biolabs Inc., Mississauga, Ontario, Canada. The Multi-Pol DNA Sequencing System was from Clontech, through BIO/CAN Scientific, Mississauga, Ontario, Canada. Fetal bovine serum (FBS) was from Bocknek, Rexdale, Ontario, Canada. The HA Test 50 was from Pharmacia, Montreal, Quebec, Canada. Recombinant N-glycosidase F was from Boehringer Mannheim, Laval, Quebec, Canada. NitroBlue Tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate were from Sigma, St. Louis, MO, U.S.A. The goat anti-(mouse IgG2b) alkaline phosphatase conjugate was from Caltag, and the modified and linearized *Autographica californica* nuclear polyhedrosis virus (AcNPV) DNA was from Pharmingen, both through Cedarlane Laboratories, Hornby, Ontario, Canada. The mouse anti-(rat chondrosarcoma link protein 2) monoclonal antibody 9/30/8-A-4 was obtained from the Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA, U.S.A.

Construction of recombinant plasmids

The complete coding sequence of human link protein plus adjacent 5' and 3' untranslated sequences (nucleotides –202 to 1092, described in [5]) were amplified by the PCR with primers LP-1 and LP-2 (Table 1). Total RNA, extracted directly from cartilage, was used as template [25]. This PCR fragment was digested with restriction endonucleases *NheI* and *SpeI*, purified

on DEAE-cellulose paper, and cloned into the *XbaI* site of pUC19 to create the recombinant plasmid pL, containing nucleotides –101 to 1083 of link protein cDNA. pL was then used as the template for the preparation of constructs to be used in the baculovirus expression system (Figure 1). A variety of synthetic oligonucleotide primers (Table 1) was used to produce the various constructs using the PCR (Table 2). Primers were synthesized using an Applied Biosystems 392 DNA synthesizer. In two cases, the PCR products were digested with the restriction endonuclease *NheI* and the excised cDNA was ligated directly into the expression vector pBlueBac2. In the other two cases, where the PCR products contained the entire pUC19 sequence, the fragments were blunt-ended and recircularized with T4 DNA ligase to form the intermediate plasmids pL2 and pL4. Using pL2 or pL4 as template, a second PCR amplification was performed to produce the desired link protein cDNA sequences, which were then excised with *NheI* and inserted into pBlueBac2 as above. Each link protein cDNA was cloned into pBlueBac2 in the correct orientation so that transcription was under the control of the polyhydriin promoter. All inserts were sequenced by the dideoxy chain-termination method [26], using the Multi-Pol DNA Sequencing System, to verify that no sequence changes had been introduced by the manipulations performed to construct the respective recombinant plasmids.

Construction of recombinant baculoviruses

Sf-9 cells, derived from *Spodoptera frugiperda* ovarian cells, were maintained in culture at 27 °C in TNM-FH medium (Grace's insect cell medium, supplemented with 3.33 g/l each of Yeastolate and lactalbumin hydrolysate) containing 10% (v/v) FBS [27]. Link protein sequences were transferred to the viral genome by cotransfection using a modified calcium phosphate precipitation procedure, as described by the supplier (Baculogold Transfection Kit, Pharmingen). Sf-9 cells (2×10^6 per 25 cm² flask) were transfected with a mixture of 2 µg of recombinant plasmid DNA and 0.5 µg of modified and linearized AcNPV DNA. The precipitate was left on the cells for 4 h at 27 °C. Monolayers were then washed twice with TNM-FH medium and incubated with 3 ml TNM-FH containing 10% FBS at 27 °C. Culture medium containing recombinant virus particles was harvested 3 days after infection. Two rounds of viral amplification were then performed to obtain a high-titre stock solution. Briefly, Sf-9 cells (9×10^6 per 75 cm² flask) were seeded in 10 ml TNM-FH containing 10% FBS and allowed to stand at room temperature for 15 min to permit cells to attach. The medium was then replaced by fresh medium containing 500 µl of transfection medium and incubated at 27 °C for 1 h. The transfection medium was replaced with 10 ml TNM-FH containing 10% FBS and the cells were incubated for 4 days at 27 °C. The medium containing the amplified recombinant viral particles was again collected, and the amplification procedure was repeated.

Production of recombinant protein

Sf-9 cells in suspension (10^7 per ml) were infected for 1 h at 27 °C with recombinant virus at 20 plaque forming units per cell. Sf-900 serum-free culture medium was then added to dilute the cells to $(1.5\text{--}2.0) \times 10^6$ cells per ml. Cells were grown at 27 °C in spinner cultures for 4 days after infection. Preliminary experiments had shown that longer culture times could yield more recombinant protein, but that its functional properties would decline. Cell culture medium was centrifuged at 5000 g for 10 min to remove cells and debris, and the supernatant was stored at 4 °C. After 4 days of culture, the production of

Table 1 PCR primers used for expression vector construction

Nucleotide numbers correspond to the published human link protein sequence [5]. *NheI* sites (GCTAGC) engineered into primers are in bold print; the STOP codon (TTA) engineered into the primer LP-7 is underlined.

Primers	Sequence	Nucleotides
LP-1	5'-CAGCGCCACAAAGAGACATT	–202 to 183
LP-2*	5'-TTAAACTAGTGCCTCTAAG	1073 to 1092
LP-3	5'-ATTCTTTGG GCTAGC AAGATGAAG	–18 to 6
LP-4*	5'-TCGAC GCTAGC GCCTCTAAG	1073 to 1087†
LP-5*	5'-CTTGTAAGTCCAGTGTACCACAACA	447 to 472
LP-6	5'-GCCGTTTTTACTATCTGATCCACCCC	776 to 801
LP-7*	5'-GG GCTAGC TCAGATATTAACCGGC	776 to 801
LP-8*	5'-CTTGGATGTGAATAGCTCTGT	80 to 100
LP-9	5'-GTGTGGTATTCCCTTACTTTC	473 to 493

* Primers represent sequence in the non-coding DNA strand.

† The first six nucleotides are from the flanking pUC19 vector sequence.

Figure 1 Preparation of link protein constructs

The entire coding region of human link protein was inserted into pUC19 to yield the plasmid pL. PCR was then used to prepare a variety of derivatives encompassing the intact link protein coding region or lacking one functional domain (A, B or B'). In two cases (L1 and L3), the PCR product was digested with *NheI* and inserted into pB2 to yield the plasmids pB2L1 and pB2L3. In the other two cases (L2 and L4), where the entire pUC19 sequence is still present, the PCR product was recircularized to yield the intermediate plasmids pL2 and pL4. A second PCR product containing only the mutant link protein sequence was then generated and inserted into pB2 to yield the plasmids pB2L2 and pB2L4. The diagram also illustrates the PCR primer recognition sites (LP3-9), the start (ATG) and stop (TGA, TAA) codons for protein translation, and the location of the baculovirus polyhedrin promoter controlling translation.

recombinant protein was estimated at 1–10 µg/ml based on detection of protein by immunoblot assay.

Detection of recombinant link protein expression

Expressed link protein was detected by immunoblot assay. Aliquots of culture supernatant were electrophoresed by the method of King and Laemmli [28] using 12.5% (w/v, total concentration) acrylamide slab gels. Proteins were transferred electrophoretically from the polyacrylamide gel to a nitrocellulose sheet [29], and immune localization of link protein was performed at room temperature. The nitrocellulose sheet was blocked with 3% (w/v) BSA in Dulbecco's PBS for 1 h. It was then immersed for 2 h in mouse monoclonal antibody 9/30/8-A-4 diluted 1:1000 in PBS containing 3% BSA, followed by 3 washes of 5 min each in PBS containing 0.1% (v/v) Tween 20. The sheet was then

incubated with a goat anti-(mouse IgG2b) alkaline phosphatase conjugate for 30 min, and washed as above. The alkaline phosphatase reaction was carried out using 0.33 mg/ml Nitro-Blue Tetrazolium and 0.16 mg/ml 5-bromo-4-chloro-3-indolyl phosphate in a buffer containing 0.1 M Tris, 0.1 M NaCl, 25 µM ethanolamine, 2 mM MgCl₂·6H₂O and 1 µM ZnCl₂, pH 9.55. The reaction was stopped with water.

Treatment with N-glycosidase F

An aliquot of the recombinant link protein in culture medium was mixed with an equal volume of N-glycosidase F in 0.125 M Tris/HCl, pH 6.8 containing 0.1% (w/v) SDS and 0.5% (w/v) Nonidet P-40. The enzyme solution was prepared by diluting the commercial preparation 1:20 with the requisite buffer to yield a concentration of 1 unit per 20 µl. The link protein/enzyme

Table 2 PCR annealing conditions

For annealing, all prior denaturation was performed at 94 °C for 45 s, and all subsequent extension was at 72 °C for 2 min.

Primer 1	Primer 2	Cycles	Annealing	
			Temperature (°C)	Time (s)
LP-1	LP-2	25	58	30
LP-3	LP-4	5*	37	30
		25	60	30
LP-5	LP-6	25	60	45
LP-3	LP-7	5*	37	30
		25	60	30
LP-8	LP-9	25	58	45

* For reactions in which primers contain engineered mismatched bases, an initial 5 cycles of the PCR were performed at 37 °C to maximize hybridization.

mixture was incubated overnight at 37 °C, before analysis by SDS/PAGE and immunoblotting as described above.

Radiometric hyaluronate assay

A modification of the Pharmacia HA test radiometric assay was used to verify that no hyaluronate or hyaluronate-binding proteins were being produced by uninfected Sf-9 cells or Sf-9 cells infected by baculovirus containing an empty expression vector. Increasing concentrations of [¹²⁵I]HABP solution (HABP, hyaluronate-binding protein), a combination of bovine link protein and the hyaluronate-binding region of aggrecan, were incubated for 1 h with fixed volumes of hyaluronate coupled to Sepharose beads. Samples were centrifuged for 10 min at 1500 *g*, the supernatants were decanted off, and the pellet was counted in a γ -radiation counter for 1 min per sample. Optimal conditions for competitive binding experiments were determined as 200 μ l of [¹²⁵I]HABP and 20 μ l of hyaluronate–Sepharose. For analysis of products secreted by Sf-9 cells, 1 ml medium, from uninfected cultures or from cultures infected with recombinant baculovirus, was added to 200 μ l of [¹²⁵I]HABP. 20 μ l of hyaluronate–Sepharose was then added, and samples were incubated for 1 h then processed as above. All manipulations were performed at room temperature.

Hyaluronate binding assay

Culture supernatants containing similar amounts of protein, based on immunoblot staining with the antibody 9/30/8-A-4, were incubated rotating for 1 h at room temperature with hyaluronate coupled to Sepharose beads. The beads were pelleted for 5 s in a microcentrifuge and the supernatant containing the unbound material was recovered. The beads were washed 3 times in 50 mM Tris, 150 mM NaCl, pH 7.5. The bound fraction was eluted by boiling the beads for 3 min in SDS/PAGE loading buffer. Both bound and unbound fractions were evaluated by gel electrophoresis and immunoblotting as described above. The hyaluronate–Sepharose beads were prepared by the method of Tengblad [30], using oligosaccharides derived by partial digestion of hyaluronate by testicular hyaluronidase. To check the speci-

ficity of interaction, an additional experiment was performed, in which the hyaluronate–Sepharose beads were mixed with increasing amounts of soluble hyaluronate (40–640 μ g/ml) before addition of the culture supernatant containing the recombinant link proteins. Bound link protein was then recovered as described above, and its concentration was analysed by immuno-slot blot analysis, using nitrocellulose membranes and antibody 9/30/8-A-4 for detection.

Aggrecan binding assay

Aggrecan from bovine nasal cartilage was purified by caesium chloride density gradient centrifugation [31]. Elution profiles of purified aggrecan and link protein in culture supernatant were established independently on a Sepharose CL-4B column (0.5 cm $\times$ 20 cm) equilibrated with 50 mM Tris, 150 mM NaCl, pH 6.8. Fractions (200 μ l) were eluted in the same buffer at a rate of 1 ml/min. The presence of aggrecan was determined by the Dimethylmethylene Blue assay [32]. The elution profile of link protein was established by immuno-slot blot analysis. Briefly, samples were denatured by boiling for 3 min in SDS/PAGE loading buffer without Bromophenol Blue, and then slot-blotted to nitrocellulose membrane. The presence of link protein was detected with the antibody 9/30/8-A-4 as described above. To study its interaction, 200 μ l of culture supernatant containing normal or mutant recombinant link protein was incubated with 50 μ l of aggrecan (5 mg/ml in 50 mM Tris, 150 mM NaCl, pH 6.8) for 1 h at room temperature. The mixture was applied onto the Sepharose CL-4B column, and fractions (200 μ l) were collected and monitored for the presence of link protein by immuno-slot blot analysis as above.

RESULTS

Four human link protein cDNA constructs were engineered into pBlueBac2 (pB2) from the original plasmid (pL) containing the entire link protein coding region and flanking untranslated DNA (Figure 1). The first construct, encompassing the entire coding region with its native translation initiation codon and stop codon, was prepared directly by PCR. The primers (LP-3 and LP-4, Table 1) were synthesized to include *Nhe*I recognition sequences external to the initiation and stop codons. Appropriate endonuclease treatment then enabled the PCR product to be inserted into the *Nhe*I site of pB2. The resulting plasmid, pB2L1, reflects the intact link protein. A similar protocol was used to construct a plasmid (pB2L3) containing an insert that reflects the link protein missing its carboxy terminal B' domain. In this case, the downstream primer (LP-7, Table 1) not only possessed an external *Nhe*I recognition sequence, but also contained a central stop codon. This stop codon was situated such that it followed the B domain in the PCR-amplified product. As with other domain boundaries used in this work, it was selected based on the known exon/intron boundaries in the chick link protein gene [10] (Figure 2).

The remaining two constructs required a more complex mode of preparation because they possessed internal deletions. The plasmid lacking the B domain (pB2L2) was prepared in a three-step process. PCR was used in a reverse orientation to that described above to amplify a product whose central region reflects the original vector of pL and whose terminal regions reflect the A domain/N-terminus of link protein and the B' domain of link protein, respectively. This PCR product was then blunt-ended, phosphorylated, and ligated to yield the plasmid pL2, which lacked the B domain at its natural exon/intron boundaries. Preparation of pB2L2 from pL2 then followed the

Figure 2 Structure of intact and mutant recombinant link proteins

The intact mature form of human link protein is depicted with its functional domains (A, B and B') indicated. Also indicated are the sites of disulfide bonds (●●), N-linked oligosaccharide substitution (Y), and exon/intron junctions (). The nucleotide triplets of codons spanning the splice junctions are depicted beneath the diagram together with the amino acid that they encode. Codon and amino acid variations resulting from the generation of the mutants used in this work are also indicated. Nucleotide positions are relative to the translation start site, and amino acid positions are relative to the N-terminus of the mature protein.

Figure 3 Expression of recombinant link protein

Recombinant protein present in insect cell culture medium was analysed by SDS/PAGE. The samples analysed were from cells infected with the recombinant baculovirus containing the insert from pB2L1 (lane 1), pB2L2 (lane 2), pB2L3 (lane 3) and pB2L4 (lane 4). Also analysed was medium from cells infected with the empty pB2 vector (lane 5), and medium from uninfected cells (lane 6). The position of reference molecular-mass standards is indicated.

same steps used in the preparation of pB2L1 from pL. A similar strategy was used to construct the plasmid lacking the A domain (pB2L4), except that the PCR primers used for the generation of the intermediate pL4 flanked the A domain instead of the B domain (Figure 1). In all cases, the four plasmids maintained the DNA encoding the link protein signal peptide, so that the expressed protein should retain the ability to enter the lumen of the rough endoplasmic reticulum and be processed for extracellular secretion. The expressed protein from plasmids encoding the intact link protein and mutants lacking the B or B' domains possess two consensus sequences for N-linked oligosaccharide substitution, whereas the mutant lacking the A domain has only one such site (Figure 2).

The four pB2 plasmids described above (pB2L1–pB2L4) all

Figure 4 Treatment of recombinant link protein with N-glycosidase F

Intact recombinant link protein present in insect cell culture medium from cells infected with the recombinant baculovirus containing the insert from pB2L1 was analysed by SDS/PAGE either before (lane 1) or after (lane 2) treatment with N-glycosidase F. The position of reference molecular mass standards is indicated.

generated recombinant proteins when expressed by Sf-9 cells infected with the relevant recombinant baculovirus. The recombinant proteins were collected in the culture medium and analysed directly by SDS/PAGE without any initial purification. Link protein was detected by immunoblotting using monoclonal antibody 9/30/8-A-4 (Figure 3). This antibody reacts with an epitope within both the B and B' domains of link protein [33], which is expected to be present in two copies on the recombinant molecules generated by pB2L1 (encoding the intact link protein) and pB2L4 (lacking the A domain), and in one copy on the recombinant molecules generated by pB2L2 and pB2L3 (lacking the B or B' domain, respectively). In all cases, the antibody recognized two proteins separated in molecular mass by 1–2 kDa, though the relative abundance of the two forms did vary between different preparations. For the recombinant intact link protein, the average molecular mass was ~46 kDa whereas all the recombinant mutants had average molecular masses of ~35 kDa. The two components observed in each case presumably reflect heterogeneity in N-linked oligosaccharide substitution, as all constructs retain the domain encoding the N-terminal consensus sequence for oligosaccharide attachment (Figure 2), which exhibits an incomplete substitution in the native link protein [34]. To verify this point, the recombinant intact link protein was treated with N-glycosidase F before being analysed by SDS/PAGE. This enzyme cleaves between the asparagine and N-acetylglucosamine residues that serve to link the oligosaccharides to the protein, and converted the two forms

Table 3 Analysis of culture medium for hyaluronate and HABPs

The data are given as means $\pm$ S.D., $n = 4$.

Sample	Radioactivity (c.p.m.)
Buffer	2309 $\pm$ 180
Uninfected Sf-9 cells	2348 $\pm$ 150
Cells infected with empty vector	2336 $\pm$ 107
Cells infected with pB2L1	1263 $\pm$ 39

Figure 5 Interaction of recombinant link protein with hyaluronate

Recombinant protein present in insect cell culture medium was analysed by SDS/PAGE, following incubation with hyaluronate–Sepharose. Samples represent material bound to the resin (lanes 1–4) and that which remained unbound (lanes 5–8). Samples analysed were from cells infected with the recombinant baculovirus containing the insert from pB2L1 (lanes 1 and 5), pB2L2 (lanes 2 and 6), pB2L3 (lanes 3 and 7), and pB2L4 (lanes 4 and 8). The positions of reference molecular mass standards are indicated.

Figure 6 Specificity of interaction between recombinant link protein and hyaluronate

Recombinant protein present in insect cell culture medium was analysed by slot blotting, following incubation with hyaluronate–Sepharose. Samples represent material bound to the resin in the presence of increasing concentrations of exogenous hyaluronate (0–640 µg/ml). Samples were from cells infected with the recombinant baculovirus containing the insert from pB2L1 (lane 1), pB2L2 (lane 2), pB2L3 (lane 3) and pB2L4 (lane 4).

of the intact link protein to a single form of smaller size, which has a molecular mass of ~ 42 kDa (Figure 4).

No molecule cross-reacting with the antibody was observed in culture medium derived from non-infected Sf-9 cells or those infected with the empty pB2 vector (Figure 3). Furthermore, analysis of these two control culture media, using an assay based

on the interaction of ^{125}I -labelled HABPs with hyaluronate–Sepharose, showed that they did not contain either hyaluronate or any endogenous HABPs. Neither medium showed any ability to compete with the binding of the iodinated-binding protein to the hyaluronate–Sepharose, unlike the situation with the culture medium containing the recombinant intact link protein (Table 3).

The ability of the recombinant link proteins to interact with hyaluronate was assessed by allowing the infected Sf-9 cell culture media to interact with hyaluronate oligosaccharides coupled to Sepharose beads. The beads and their retained material were separated by centrifugation, and the presence of link protein attached to the beads or remaining in the medium was assessed by SDS/PAGE and immunoblotting (Figure 5). The intact recombinant link proteins all interacted with the hyaluronate, suggesting that the baculovirus expression system allows a normal protein conformation to form with normal disulphide bond formation. All three mutant link proteins were also able to interact with hyaluronate under these conditions, and again no non-functional molecules were observed. The specificity of the interaction between the immobilized hyaluronate and the various recombinant link proteins was verified by performing competition experiments in the presence of soluble hyaluronate (Figure 6). In all cases, the presence of increasing concentrations of soluble hyaluronate served to progressively inhibit binding to the immobilized hyaluronate. Thus, a lack of the A domain does not interfere with the ability of the recombinant protein to interact with hyaluronate, and the presence of either the B or B' domain is sufficient for this function. The data also illustrate that native folding of the B and B' domains appears to occur in the baculovirus expression system, irrespective of whether the normal adjacent structural domain is present or not.

The ability of the recombinant link proteins to interact with aggrecan was also assessed. This was performed using a gel filtration system, in which free aggrecan and link protein elute near the void volume and total volume of the column, respectively. The elution of the free link protein as a discrete component at the total volume of the column suggests that self-association of link protein is not a problem at the low concentrations used in the present work, and allows interaction of link protein with aggrecan in mixtures of the two components to be indicated by accelerated elution of the link protein. The intact recombinant link protein eluted entirely at the void volume of the column in association with aggrecan (Figure 7), as one would expect if the protein has a native conformation and is fully functional. The two mutant link protein molecules possessing the A domain also showed such an interaction, with the link proteins coeluting with aggrecan. The mutant link protein lacking the A

Figure 7 Interaction of recombinant link protein with aggrecan

Recombinant protein present in insect cell culture medium was analysed by slot blotting, following fractionation by gel filtration. Samples from cells infected with the recombinant baculovirus containing the insert from pB2L1 (lane 2), pB2L2 (lane 3), pB2L3 (lane 4), and pB2L4 (lane 5) were analysed following chromatography in the presence of aggrecan. The elution position of the intact recombinant link protein in the absence of aggrecan is also indicated (lane 1), as are the chromatography fraction numbers.

domain showed little ability to interact under these conditions, and eluted predominantly near the total column volume. Thus the presence of the A domain appears to be an absolute requirement for strong interaction with aggrecan. As described above for the B and B' domains, the data suggest that native folding of the A domain has occurred in all cases, irrespective of the structure of the neighbouring C-terminal domain in the recombinant molecules.

DISCUSSION

The baculovirus expression system has been used to generate four recombinant proteins. One protein represents the full-length sequence of human link protein, and the other three represent mutants in which the protein encoded by one exon has been deleted. Two of the mutants (pB2L2 and pB2L4) were generated by the use of PCR to excise the requisite DNA from the original plasmid containing the entire link protein sequence (pL). In each case, the corresponding exon boundaries split codons, to leave a single base at their 5' end and two bases at their 3' end, so that upon ligation the reading frame is maintained (Figure 2). Furthermore, the newly spliced codon encodes a glycine residue in each case, which is the same amino acid that occurs in the intact link protein across the splice sites adjacent to the B domain. The third mutant (pB2L3) utilizes an induced stop codon in its expression, to prevent the synthesis of the C-terminal protein domain. This stop codon does, however, allow the expression of the glycine residue bridging the natural splice site and the following three amino acids (Figure 2). Thus, the mutants encoded by pB2L2, pB2L3 and pB2L4 lack amino acids 143–243 (B domain), 247–339 (B' domain) and 19–142 (A domain) respectively of the mature intact human link protein.

It appears that folding of the protein occurs normally in this insect cell culture system, as all expressed domains show their expected function. This presumably indicates normal disulphide bond formation and normal conformation within each domain. Certainly, conformational changes induced by reduction and alkylation of disulphide bonds are known to prevent the participation of the native link protein in proteoglycan aggregate formation [35]. Thus, it appears that the baculovirus expression system permits both secretion of recombinant proteins into the culture medium and the occurrence of native intramolecular interactions. Protein folding is always a concern when one is expressing deletion mutants, as one cannot assume *a priori* that the remaining protein sequences will adopt the same conformation as in the intact protein. In the case of link protein, the functional domains occur entirely within discrete exons and all disulphide bonds lie within the discrete domains. Thus all the information for correct protein folding of a given domain might be encoded within the corresponding exon. It was for this reason that mutants were made to delete domains at their natural exon/intron boundaries, and it would appear that this logic was valid.

Two of the mutant link proteins possess both possible sites for N-linked oligosaccharide substitution that are present in the intact link protein, whereas the mutant lacking the A domain has only the site at asparagine residue 6 (Figure 2). The second site at asparagine residue 41 falls within the exon encoding the A domain. All the mutants and the intact recombinant link protein show two components upon SDS/PAGE analysis, albeit to different extents. In the case of the intact recombinant molecule, this size heterogeneity was shown to be caused by variation in the substitution by N-linked oligosaccharides. It is thus likely that all the recombinant molecules exhibit the same variability in oligosaccharide substitution at asparagine residue 6 as the native link

protein extracted from articular cartilage [11]. The predicted molecular mass for the intact recombinant link protein is 38218 Da, and those for the mutants lacking the A, B or B' domain are 24415, 26664 and 27612 Da respectively. For the intact link protein and the mutants lacking the B or B' domain, the observed size upon SDS/PAGE is ~8 kDa larger than that predicted, suggesting a similar size for the N-linked oligosaccharides. For the mutant lacking the A domain, the discrepancy in size is ~10 kDa. This is larger than expected for a molecule which has only one site for oligosaccharide substitution. It is not clear whether this size variation is an artefact of SDS/PAGE or truly reflects a larger oligosaccharide. As the size of the intact recombinant link protein on SDS/PAGE is comparable to that observed for the native link protein extracted from cartilage, it is apparent that the insect cells are capable of synthesizing N-linked oligosaccharides of a similar size to those produced by the human.

It is apparent from this work that the A domain is essential for the interaction between the link protein and aggrecan, and that the B and B' domains do not contribute substantially to this interaction. Such a situation had been predicted from studies using proteolysis to fragment the two interacting molecules [22]. It is also apparent that both the B and B' domains are able to interact independently with hyaluronate, and that absence of the A domain does not interfere with this interaction. Again, proteolytic fragmentation studies had indicated that the B and B' domains were involved in this interaction [22], though it was not clear from this work whether either domain could interact alone. In addition to link protein, the proteoglycans aggrecan [7], versican [36] and neurocan [37] also possess the ability to interact with hyaluronate, and in all cases the region responsible for this interaction shares sequence identity with both the B and B' domains. These domains have been termed proteoglycan tandem repeats [38]. Not all such domains possess the ability to interact with hyaluronate, however. Aggrecan possesses two globular regions at the N-terminus of its core protein. The external region, termed the G1 region, has areas that share sequence identity with the A, B and B' domains of link protein and is functionally active [8]. The adjacent G2 region has areas that share sequence identity with the B and B' domains, yet shows no interaction with hyaluronate [39]. In addition, not all proteins that interact with hyaluronate possess both B and B' domains. The lymphocyte homing receptor, CD44, has a region that exhibits sequence identity with only a single B-like domain, yet is able to interact with hyaluronate [40,41]. Such an interaction is analogous to that exhibited by two of the mutant link proteins expressed in this work.

It remains to be established whether molecules possessing two functional B-type domains interact with greater affinity than those possessing only a single domain. Moreover, in the case of the mutant link proteins, it must still be demonstrated whether the absence of a B-type domain impairs the ability of the molecule to participate in proteoglycan aggregate formation or stability. If this participation is not impaired, then naturally occurring gene defects affecting splicing that result in deletion of a B-type domain would not be expected to cause a problem with cartilage function. It would also be predicted that single base substitutions causing amino acid changes that affect the interactive properties of one B-type domain would also not result in impaired cartilage function. This may not be the case with natural mutations affecting the single link protein A domain.

We would like to thank the Shriners of North America and the National Institutes of Health for their financial support, Jane Wishart for preparing the figures, and Nadia Nikolajew for typing the manuscript.

REFERENCES

- 1 Hascall, V. C. (1977) *J. Supramol. Struct.* **7**, 101–120
- 2 Poole, A. R., Pidoux, I., Reiner, A. and Rosenberg, L. (1982) *J. Cell Biol.* **92**, 921–937
- 3 Hardingham, T. E. (1979) *Biochem. J.* **177**, 237–247
- 4 Osborne-Lawrence, S. L., Sinclair, A. K., Hicks, R. C., Lacey, S. W., Eddy, R. L., Byers, M. G., Shows, T. B. and Duby, A. D. (1990) *Genomics* **8**, 562–567
- 5 Dudhia, J. and Hardingham, T. E. (1990) *Nucleic Acids Res.* **18**, 1292
- 6 Neame, P. J., Christner, J. E. and Baker, J. R. (1986) *J. Biol. Chem.* **261**, 3519–3535
- 7 Doege, K. J., Sasaki, M., Kimura, T. and Yamada, Y. (1991) *J. Biol. Chem.* **266**, 894–902
- 8 Paulsson, M., Mörgelin, M., Widemann, H., Beardmore-Gray, M., Dunham, D., Hardingham, T., Heinegård, D., Timpl, R. and Engel, J. (1987) *Biochem. J.* **245**, 763–772
- 9 Deak, F., Kiss, I., Sparks, K. J., Argraves, W. S., Hampikian, G. and Goetinck, P. F. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3766–3770
- 10 Kiss, I., Deak, F., Mestric, S., Delius, H., Soos, J., Dékány, K., Argraves, S. W., Sparks, K. S. and Goetinck, P. F. (1987) *Proc. Natl. Acad. Sci.* **84**, 6399–6403
- 11 Liu, J., Cassidy, J. D., Allan, A., Neame, P. J., Mort, J. S. and Roughley, P. J. (1992) *J. Orthop. Res.* **10**, 621–630
- 12 Mort, J. S., Poole, A. R. and Roughley, P. J. (1983) *Biochem. J.* **214**, 269–272
- 13 Nguyen, Q., Liu, J., Roughley, P. J. and Mort, J. S. (1991) *Biochem. J.* **278**, 143–147
- 14 Oegema, T. R., Brown, M. and Dziewiatkowski, D. (1977) *J. Biol. Chem.* **252**, 6470–6477
- 15 Faltz, L. L., Caputo, C. B., Kimura, J. H., Schrode, J. and Hascall, V. C. (1979) *J. Biol. Chem.* **254**, 1381–1387
- 16 Mörgelin, M., Paulsson, M., Hardingham, T. E., Heinegård, D. and Engel, J. (1988) *Biochem. J.* **253**, 175–185
- 17 Caterson, B. and Baker, J. (1978) *Biochem. Biophys. Res. Commun.* **80**, 496–503
- 18 Franzén, A., Björnsson, S. and Heinegård, D. (1981) *Biochem. J.* **197**, 669–674
- 19 Rosenberg, L., Tang, L.-H., Pal, S., Johnson, T. L. and Choi, H. (1988) *J. Biol. Chem.* **263**, 18071–18077
- 20 Bonnet, F., Dunham, D. G. and Hardingham, T. E. (1985) *Biochem. J.* **228**, 77–85
- 21 Kimata, K., Hascall, V. C. and Kimura, J. H. (1982) *J. Biol. Chem.* **257**, 3827–3832
- 22 Périn, J.-P., Bonnet, F., Thuriereau, C. and Jollès, P. (1987) *J. Biol. Chem.* **262**, 13269–13272
- 23 Goetinck, P. F., Stirpe, N. S., Tsonis, P. A. and Carlone, D. (1987) *J. Cell Biol.* **105**, 2403–2408
- 24 Neame, P. J., Christner, J. E. and Baker, J. R. (1987) *J. Biol. Chem.* **262**, 17768–17778
- 25 Grover, J. and Roughley, P. J. (1993) *Biochem. J.* **291**, 361–367
- 26 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- 27 Summers, M. D. and Smith, G. E. (1988) *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures* (Texas Agricultural Experiment Station Bulletin No. 1555), pp. 38–42, Texas A&M University, College Station
- 28 King, J. and Laemmli, U. K. (1971) *J. Mol. Biol.* **62**, 465–477
- 29 Towbin, M., Staehelin, T. and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4350–4354
- 30 Tengblad, A. (1979) *Biochim. Biophys. Acta* **578**, 281–289
- 31 Santer, V., White, R. J. and Roughley, P. J. (1981) *Arthritis Rheum.* **24**, 691–700
- 32 Farndale, R. W., Sayers, C. A. and Barrett, A. J. (1982) *Connect. Tissue Res.* **9**, 247–248
- 33 Hughes, C. E., Caterson, B., Calebro, A. and Hascall, V. C. (1993) *Trans. Orthop. Res. Soc.* **18**, 676
- 34 Nguyen, Q., Murphy, G., Roughley, P. J. and Mort, J. S. (1989) *Biochem. J.* **259**, 61–67
- 35 Hascall, V. C. and Sajdera, S. W. (1969) *J. Biol. Chem.* **244**, 2384–2396
- 36 Zimmermann, D. R. and Ruoslahti, E. (1989) *EMBO J.* **8**, 2975–2981
- 37 Rauch, U., Karthikeyan, L., Maurel, P., Margolis, R. U. and Margolis, R. K. (1992) *J. Biol. Chem.* **267**, 19536–19547
- 38 Hardingham, T. E. and Fosang, A. J. (1992) *FASEB J.* **6**, 861–870
- 39 Fosang, A. J. and Hardingham, T. E. (1989) *Biochem. J.* **261**, 801–809
- 40 Goldstein, L. A., Zhou, D. F. H., Picker, L. J., Minty, C. D., Bargatze, R. F., Ding, J. F. and Butcher, E. C. (1989) *Cell* **56**, 1063–1072
- 41 He, Q., Lesley, J., Hyman, R., Ishihara, K. and Kincade, P. W. (1992) *J. Cell Biol.* **119**, 1711–1719