Roles of aggrecan domains in biosynthesis, modification by glycosaminoglycans and product secretion

Chris KIANI, Vivian LEE, Liu CAO, Liwen CHEN, Yaojiong WU, Yaou ZHANG, Mark E. ADAMS and Burton B. YANG[1]

Sunnybrook & Women's College Health Sciences Centre, 2075 Bayview Avenue, Toronto M4N 3M5, Canada, and Department of Laboratory Medicine and Pathobiology, University of Toronto, 100 College Street, Toronto M5G 1L5, Canada

Aggrecan is a member of the chondroitin sulphate (CS) proteoglycan family, which also includes versican/PG-M, neurocan and brevican. Members of this family exhibit structural similarity: a G1 domain at the N-terminus and a G3 domain at the C-terminus, with a central sequence for modification by CS chains. A unique feature of aggrecan is the insertion of three additional domains, an inter-globular domain (IGD), a G2 domain and a keratan sulphate (KS) domain (sequence modified by KS chains), between the G1 domain and the CS domain (sequence modified by CS chains). The G1 and G3 domains have been implicated in product secretion, but G2, although structurally similar to the tandem repeats of G1, performs an unknown function. To define the functions of each aggrecan domain in product processing, we cloned and expressed these domains in various combinations in COS-7 cells. The results indicated that the G3 domain enhanced product secretion, alone or in combination with the KS or CS domain, and promoted glycosaminoglycan (GAG) chain attachment. Constructs containing the G1 domain were not secreted. Addition of a CS domain sequence to G1 reduced this inhibition, but GAG chain attachment was still decreased. The potential GAG chain attachment site in the IGD was occupied by GAGs, and IGD product was secreted efficiently. The KS domain was modified by GAG chains and secreted. Finally, the G2 domain was expressed but not secreted, and inhibited secretion of the IGD when expressed as an IGD–G2 combination.

Key words: chondroitin sulphate, G1 domain, G2 domain, G3 domain, glycosaminoglycan, keratan sulphate.

INTRODUCTION

The proteoglycans are a family of glycoconjugates with a central core protein to which glycosaminoglycan (GAG) side chain(s) are covalently linked post-translationally [1]. The majority of the functions of proteoglycans are mediated by these GAG chains [2,3], which are polymers of repeated dissacharide units consisting of a uronic acid and a hexosamine. Biosynthesis of all GAGs, except hyaluronic acid, is initiated from a core protein. GAGs are O-linked to serine residues through a trisaccharide linkage sequence at their reducing ends: serine–xylose–galactose–galactose. These GAG chains are acidic molecules, and they participate in a wide variety of binding interactions with other matrix macromolecules, cations and water [4,5]. They can concentrate secretory products surrounding cells, and sequester a variety of extracellular proteins at cell surfaces. In cartilage, the matrix molecules that make up the extracellular matrix (ECM) include proteoglycans, hyaluronic acid, type II collagen, glycoproteins and various mixtures of elastic fibres. Aggrecan is the major structural proteoglycan in cartilage and is responsible for cartilage's resilience and load-bearing properties. Loss of aggrecan is a major feature of cartilage degradation associated with arthritis [6,7].

The core protein of aggrecan is composed of three globular domains (G1, G2 and G3), with one inter-globular domain (IGD) linking G1 and G2, and two exons for keratan sulphate (KS) chain attachment (KS domain) and for chondroitin sulphate (CS) chain attachment (CS domain) situated between G2 and G3. Attachment of these GAG chains occurs on the serine of serine–glycine dipeptide sequence present in this region, and one molecule of aggrecan can contain up to 100 CS chains, 30 KS chains and many O- and N-linked oligosaccharides [1]. The G1 domain comprises the N-terminus of the core protein. This domain has the same structural motifs as link protein [8]. The G2 domain is homologous to the tandem repeats found in G1 and link protein. The G3 domain, which makes up the C-terminus of the core protein, is composed of alternatively spliced epidermal growth factor-like domains, a carbohydrate recognition domain, a complement-binding-protein-like domain and a short tail [9,10]. Recent studies indicated that the G1 domain is poorly secreted, and that the G3 domain plays a role in the secretion of recombinant products [11]. However, the effects of the other domains on product secretion are not known.

Given that there is up to 62.1 % identity between G1 and G2 tandem repeat motifs [2], we reasoned that G2 might also play a role in product secretion. To test this, we have performed a comprehensive study of the role of the six domains of aggrecan in product processing using recombination and gene expression techniques. Our results indicated that the promotion of product secretion by G3 was inhibited by G1 and G2. This inhibition was overcome in the presence of G3 and the CS domain (i.e. sequence modified by CS chains), since G3 stimulated GAG chain attachment to this domain, and the GAG-modified product was effectively secreted. The IGD was modified by GAG chains, and the product was also well secreted.

EXPERIMENTAL

Materials

Reverse transcription–PCR mRNA amplification kit, *Taq* DNA polymerase, T4 DNA ligase and restriction endonucleases were

Abbreviations used: CS, chondroitin sulphate; CS domain, sequence modified by CS chains; DMEM, Dulbecco's modified Eagle's medium; ECM, extracellular matrix; FBS, fetal bovine serum; GAG, glycosaminoglycan; IGD, inter-globular domain; KS, keratan sulphate; KS domain, sequence modified by KS chains; LB, Luria broth.

[1] To whom correspondence should be addressed, at Sunnybrook & Women's College Health Sciences Centre (e-mail Burton.Yang@swchsc.on.ca).

Table 1 Primers and restriction endonuclease sites for PCRs

Underlined sequences represent restriction enzyme sites denoted in the primer names.

Primer	Sequence	Restriction enzyme
G1N	5′ AAA CTC GAG CTG GGA AGC TCC CTG AAC 3′	*Xho*I
G1C	5′ AAA CGG GAT CCC ATC ACC ACT GTA ACA GAT GGC ATC GTA 3′	*Bam*HI
G1C	5′ AAA GTC GAC ATC ACC ACT GTA ACA GAT 3′	*Sal*I
G1C	5′ AAA GCA TGC ATC ACC ACT GTA ACA GAT 3′	*Sph*I
IGDN	5′ AAA CTC GAG GAC TTC GAG GCT CTG GTC CCA 3′	*Xho*I
IGDR	5′ AAA TCT AGA GGG GGA AAT GGG CTG 3′	*Xba*I
G2MF	5′ AAA GCA TGC GTA GGA GAC AAA GAG AGC 3′	*Sph*I
G2MR	5′ AAA GCA TGC GCT GCG GGG ATT CAC AAT 3′	*Sph*I
G2F	5′ AAA CTC GAG GGT GTG GTG TTC CAC TAC 3′	*Xho*I
G2R	5′ AAA TCT AGA TCT GAA GCA GAA AGC GTG 3′	*Xba*I
KSN	5′ AAA GTC GAC GCT CTG CCA TCC GTA GTG 3′	*Sal*I
KSC	5′ AAA GCA TGC TGA GAG CAG TGA TAC GTC 3′	*Sph*I
KSC	5′ AAA TCT AGA TGA GAG CAG TGA TAC GTC 3′	*Xba*I
CSN	5′ AAA AAA GGA TCC ATC CCG TAT TTC AGC GGA GAC 3′	*Bam*HI
CSN	5′ AAA AAA CTC GAG ATC CCG TAT TTC AGC GGA GAC 3′	*Xho*I
CSC	5′ AAA GAA TTC AGA AGT GGC AGC AGT GTC AGT GCT 3′	*Eco*RI
CSC	5′ AAA AAA TCT AGA AGT GGC AGC AGT GTC AGT GCT 3′	*Xba*I
CSC	5′ AAA AAA CTC GAG AGT GGC AGC AGT GTC AGT GCT 3′	*Xho*I
G3N	5′ AAA GAA TTC AGT GGT GAG CCC TCC GGT GCT 3′	*Eco*RI
G3N	5′ AAA CTC GAG AGT GGT GAG CCC TCC GGT GCT CCT GAG 3′	*Xho*I
G3C	5′ AAA GCA TGC CGC CCC GCT CTA ATG GGT G 3′	*Sph*I
G3C	5′ AAA CTC GAG AGA TCC TCT CAT GCA TGC ATG GGT GGG TCT GTG CAC 3′	*Xho*I
G3CHIS	5′ AAA TCT AGA GTG ATG GTA ATG GTG ATG ATG GGT GGG TCT GTG CAC 3′	*Xba*I
LPN	5′ AAA AAA GAA TTC CTA AGT CTA CTC TTT CTG GTG CTG 3′	*Eco*RI

purchased from Boehringer Mannheim. Mammalian expression vectors (pcDNA1, pcDNA3 and pCR3.1), Unidirectional Mammalian TA Cloning Kit and *Escherichia coli* strains MC1061 and TOP10F′ were purchased from Invitrogen. Bacterial growth medium was from Difco. Prep-A-Gene DNA purification kit and prestained protein markers were from Bio-Rad. DNA mini-prep kit was from Bio/Can Scientific. Lipofectin, Dulbecco's modified Eagle's medium (DMEM), fetal bovine serum (FBS) and trypsin/EDTA were from GIBCO BRL. ECL Western Blot Detection kit was from Amersham. Horseradish peroxidase-conjugated goat anti-(mouse IgG) was from Sigma. DNA Midi-prep kit was from Qiagen Inc. Tissue culture plates (6-well and 100 mm) were from Nunc Inc. All other chemicals were from Sigma.

Strategy for cloning and construction of recombinant genes

In order to generate expression constructs containing various regions of mature aggrecan, we used the following schemes, using chicken cDNA (synthesized from chicken mRNA) as template and two primer sets, IGDN*Xho*I/G2MR*Sph*I and KSC*Xba*I/G2MF*Sph*I in PCR reactions. Sequences for all primers used in this study are shown in Table 1 according to the published sequence [12,13]. IGD–G2 (the N-terminal half of G2) was amplified using the first primer set, and G2–KS (the C-terminal half of G2) was amplified using the second primer set in PCR. These PCR products were agarose-gel-purified and doubly digested with their respective restriction enzymes. The restricted PCR fragments were further gel-purified and ligated into *Xho*I/*Xba*I-restricted mammalian expression vector pcDNA3 carrying link protein leading peptide, which is recognized by the monoclonal antibody 4B6. The cDNA of link protein leading peptide was synthesized using LPN*Eco*RI and LPC*Bam*HI as primers and the link protein cDNA as template in a PCR reaction. The PCR products were then gel-purified and doubly digested and ligated into *Eco*RI/*Bam*HI-restricted pcDNA3. The ligation mixture was then used as template to PCR-amplify the construct IGD–G2–KS using the primers LPN*Eco*RI/KSC*Xba*I utilizing the Invitrogen Unidirectional TA Cloning Kit. A 2 μl portion of the ligation mix was used to transform *E. coli* TOP10F′ cells, as described by the manufacturer. The transformation mixture was plated on a Luria broth (LB) plate containing 50 mg/ml kanamycin. The resultant construct, IGD–G2–KS, was confirmed using restriction digestion and sequencing.

IGD–G2–KS was then used as template to amplify IGD, G2, KS domain and IGD–G2 fragments using the following primer sets: IGDN*Xho*I/IGDR*Xba*I for IGD, G2F*Xho*I/G2R*Xba*I for G2, KSN*Sal*I/KSC*Xba*I for the KS domain and IGDN*Xho*I/G2R*Xba*I for IGD–G2. The PCR products were gel-purified and doubly digested overnight at 37 °C with the respective enzymes. The digested fragments were further gel-purified and used in ligation reactions (16 °C overnight) with *Xho*I/*Xba*I-restricted pcDNA3 harbouring a 5′ LP60 sequence. In order to increase the cloning efficiency, the ligation reaction mixtures were used further as templates to amplify IGD, G2, the KS domain and IGD–G2 using the following primer sets: LPN*Eco*RI/IGDR*Xba*I for IGD, LPN*Eco*RI/G2R*Xba*I for G2, LPN*Eco*RI/KSC*Xba*I for KS and LPN*Eco*RI/G2R*Xba*I for IGD–G2. The amplified products were cloned using an Invitrogen Unidirectional TA Cloning Kit.

To generate the expression construct G1–KS, we first PCR-amplified the G1 fragment from a recombinant construct containing G1, which was cloned and expressed by us previously [14–16], using the primer set LPN*Eco*RI/AG1C*Sal*I. The KS domain was also amplified, using the primers KSN*Sal*I/KSC*Xba*I, and with IGD–G2–KS as a template. This PCR product was gel-purified and doubly digested overnight with *Sal*I and *Xba*I. The digested materials were again gel-purified and ligated into *Eco*RI/*Xba*I-digested pcDNA3. The ligation mixture was used further as template to PCR-amplify the G1–KS fragment using the primer set LPN*Eco*RI/KSC*Xba*I

and cloned utilizing the Invitrogen Unidirectional TA Cloning Kit. The identity of the resultant recombinant DNA was confirmed by restriction digests and sequencing.

To generate G3–KS, we PCR-amplified the G3 moiety, which was originally cloned by Stirpe et al. [17], from the template G3 cDNA in pcDNA3 using the primers LPN*Eco*RI/G3C*Xho*I. The KS domain was also PCR-amplified from the IGD–KS construct using the primers KSN*Sal*I/KSC*Sph*I. The PCR products were gel purified and doubly digested with *Eco*RI and *Xho*I or *Sal*I and *Sph*I respectively. The digested products were purified and eventually ligated into *Eco*RI/*Sph*I-restricted pcDNA1. The ligation mixture was used in the transformation of *E. coli* strain MC1061, and the transformed cells were plated on to LB plates containing 50 μg/ml ampicillin and 10 μg/ml tetracycline.

The construction of expression plasmid G1–G2 was accomplished as follows. The G1 moiety was PCR-amplified from the G1 construct using the primers LPN*Eco*RI/G1C*Sal*I and doubly digested with *Eco*RI and *Sal*I. The G2 moiety was isolated from the *Xho*I/*Xba*I-digested G2 construct. These fragments were ligated into the *Eco*RI/*Xba*I-digested pcDNA3. The ligation mixture was used to PCR-amplify the whole fragment, G1–G2, using the primer set LPN*Eco*RI/G2R*Xba*I. The PCR products were cloned directly into pCR3.1.

To generate the G1–G3 construct, the primers LPN*Eco*RI/G1C*Sal*I were used to amplify the G1 fragment. G1 was ligated (after double digestion with *Eco*RI and *Sal*I and gel purification) into the *Eco*RI/*Xho*I-digested G3 construct (in pcDNA3). The identity of resultant recombinant, G1–G3, was further confirmed by restriction digestion and sequencing.

To construct another set of recombinant DNA fragments, i.e. the CS domain, G1–CS and CS–G3, our strategy was to PCR-amplify the following DNA fragments: link protein leading sequence, G1, CS domain and G3. The DNA fragments were then ligated together in various combinations in pcDNA3, and their identities were confirmed by restriction digestion and sequencing. Briefly, an *Eco*RI site and a *Bam*HI site were created at the 5′ and 3′ ends of link protein that was cloned previously [8]. After using the primer set CSN*Bam*HI/CSC*Xba*I and the chicken genomic DNA as template, CS domain PCR products were purified and digested with *Bam*HI and *Xba*I. The digested link protein leading sequence and CS fragment were ligated into *Eco*RI/*Xba*I-digested pcDNA3. After an overnight incubation of the ligation mixture at 16 °C, *E. coli* strain TOP10F′ was transformed with the mixture.

In order to generate CS–G3, we PCR-amplified the G3 domain using the primer set G3N*Xho*I/G3C*Xba*I, and amplified the CS fragment from the CS construct using primers LPN*Eco*RI/CSC*Xho*I. The PCR fragments were digested and ligated into *Eco*RI/*Xba*I-digested pcDNA3. The ligation mixture was used in the transformation of TOP10F′ cells.

The G1–CS construct was generated in the same manner. Using the primer sets LPN*Eco*RI/G1C*Sal*I and CSN*Xho*I/CSC*Xba*I, we amplified the G1 and CS fragments respectively. These PCR products were digested and ligated into *Eco*RI/*Xba*I-restricted pcDNA3, followed by transformation of *E. coli* strain TOP10F′.

DNA amplification, purification, ligation and transformation

DNA was amplified in a PCR reaction using pairs of appropriate primers. The reaction mixture (total final volume of 100 μl) contained 200 μM dNTPs, 0.2 μg of each primer, 50 ng of template DNA, 2 units of *Taq* DNA polymerase and Mg-containing buffer (Boehringer Mannheim). The reactions were carried out at 94 °C for 5 min for one cycle, then for 25 cycles at 94 °C for 60 s, 55 °C for 60 s and 72 °C for 60–120 s (depending on the size of DNA amplified), with a final extension at 72 °C for 10 min.

The DNA products from PCR reactions were purified using a Prep-A-Gene DNA purification kit, following the manufacturer's instructions. The purified DNA was doubly digested with two appropriate restriction endonucleases, purified and eluted into 40 μl of water. The DNA was then ligated into the appropriate plasmid (pcDNA1, pcDNA3 or pCR3.1), which had been linearized with appropriate restriction enzymes. A ligation mixture typically contained 1 μl of ligation buffer, 1 μl of DNA ligase, 3 μl of plasmid vector (50 ng) and 5 μl of insert (150 ng). The ligation reaction was carried out at 14 °C (for pCR3.1) or 16 °C (for pcDNA1 and pcDNA3) overnight. A 2 μl portion of the ligation mixture were used to transform competent *E. coli* strain MC1061 (for pcDNA1 vector backbone) or TOP10F′ (for pcDNA3 and pCR3.1 vector backbones).

To prepare electro-competent bacteria, *E. coli* cells were grown in 1 litre of LB until the density (D_{590}) reached 0.8. The cells were pelleted for 10 min at 10000 ***g***, washed twice with water, resuspended in 4 ml of water containing 10 % (v/v) glycerol, divided into aliquots (100 μl each) and stored at −70 °C. In a transformation reaction, 2–5 μl of DNA ligation mixture was combined with 100 μl of competent bacteria in a 0.2 cm cuvette and electroporated at 2.5 kV in a Bio-Rad electroporator. The settings for the gene pulser and capacitance extender were: capacitance set at 25 μF, capacitance extender set at 960 μF and resistance set at 200 Ω. The mixture was transferred to 0.5 ml of SOC medium [18], agitated at 230 rev./min for 45 min at 37 °C and spread on to regular LB agar plates containing appropriate antibiotics and cultured at 37 °C overnight.

Expression of recombinant constructs in COS cells

COS-7 cells were transiently transfected with recombinant constructs using Lipofectin (GIBCO) according to the manufacturer's instructions. Briefly, the cultured COS-7 cells were seeded on to a six-well plate (1.5×10^5 cells/well). The cells were allowed to attach and grow overnight in DMEM supplemented with 5 % (v/v) FBS. The following day the COS-7 cells had reached 70 % confluence, at which time they were ready for transfection. Lipofectin (10 μl) was incubated with plasmid DNA (5 μg) for 15 min in 200 μl of DMEM, followed by addition of 800 μl of DMEM. During the incubation, the COS-7 cell culture was rinsed with 2 ml of DMEM. The Lipofectin/DNA mixture was applied to the rinsed cultures and incubated for 10 h. Then the DNA/Lipofectin mixture was replaced with 1 ml of DMEM supplemented with 5 % (v/v) FBS. At 3 days after transfection, growth medium and cell lysate were harvested separately, and samples were frozen until analysis. For time course analysis of product expression, transfection of COS-7 cells with G1, G2 or G3 constructs was carried out as above. Culture medium and cell lysate were harvested 1, 2 and 3 days after the addition to the cultures of DMEM supplemented with 5 % (v/v) FBS, and expression of these constructs and secretion of the products was analysed by Western blotting.

Western blot assays

Cell lysate and culture medium were subjected to SDS/PAGE, as described previously [19,20]. Proteins separated by SDS/PAGE were transblotted on to a nitrocellulose membrane (Bio-

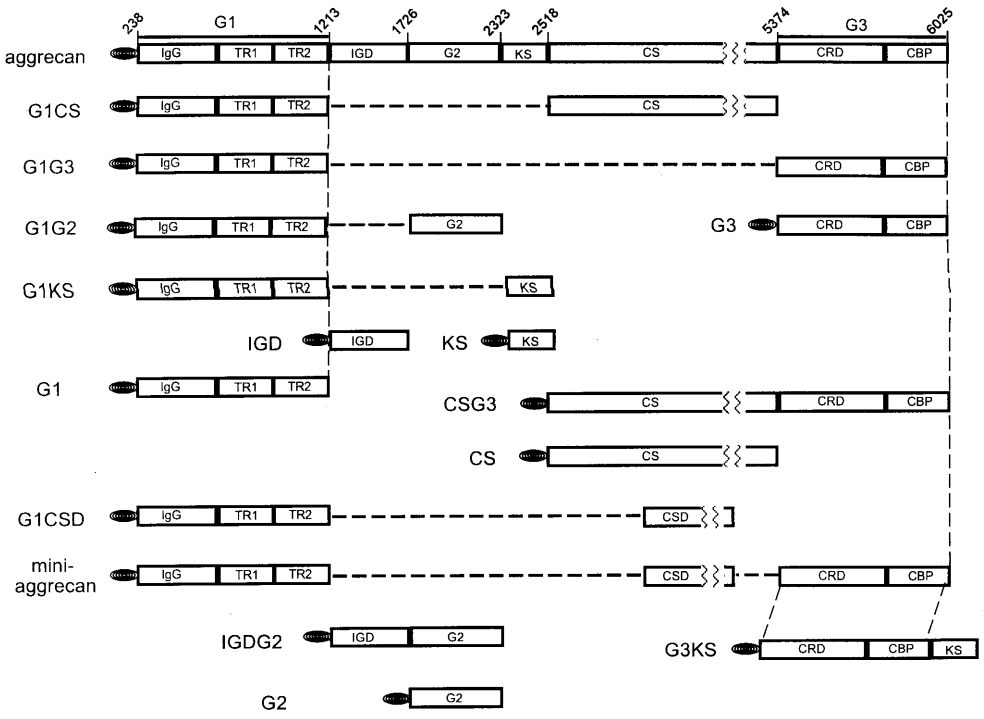

Figure 1 Strategy for the construction of 14 recombinant constructs containing different domains of aggrecan

The G1 domain includes nucleotides 144–1038 of mature aggrecan; the IGD, nucleotides 1041–1554; the G2 domain, nucleotides 1557–2145; the KS domain, nucleotides 2154–2409; the truncated CS domain, nucleotides 3795–5226; the G3 domain, nucleotides 5679–6327. Abbreviations: IgG, immunoglobulin-like motif; TR, tandem repeat; CRD, carbohydrate recognition domain; CBP, complement-binding protein domain. The leading peptide added to all constructs was obtained from link protein (nucleotides 1–180). Numbers above the diagrams correspond to nucleotides in the sequence of full-length aggrecan.

Rad) in 1 × TG buffer (Amresco) containing 20 % (v/v) methanol at 60 V for 2 h in a cold-room. The membrane was blocked in TBST (10 mM Tris/HCl, pH 8.0, 150 mM NaCl, 0.05 % Tween 20) containing 10 % non-fat dry milk powder (TBSTM) for 1 h at room temperature, and then incubated at 4 °C overnight with monoclonal antibody 4B6 diluted in TBSTM. The membranes were washed with TBST (3 × 30 min) and then incubated for 1 h with horseradish peroxidase-conjugated goat anti-mouse antibody diluted (1:1000) in TBSTM. After washing as above, the bound antibody was visualized by chemiluminescence (ECL kit; Amersham).

Treatment with chondroitinase ABC

Protein A beads (50 μl of gel slurry) were incubated with an excess amount of 4B6 antibody at room temperature for 2 h. The unbound antibody was recovered, and the gel beads were washed extensively with PBS three times. Culture medium from COS-7 cells transfected with mini-aggrecan construct was mixed with 2 × PBS in a 1:1 (v/v) ratio, followed by incubation with the antibody-bound gel beads at 4 °C overnight. The gel beads were washed extensively and resuspended in 20 μl of 1 × PBS. Chondroitinase ABC (0.25 unit) was added to digested GAG chains at 37 °C for 2 h. The digested product was recovered with 1 × protein loading dye, and analysed on Western blot probed with antibody 4B6 as above.

Densitometer analysis

Relative protein concentrations were estimated using a densitometer (Molecular Dynamics) to scan the densities of the signals after Western blotting. The relative intensity of each band after Western blot development is shown below each blot (× 10^3).

RESULTS

Expression of recombinant constructs

To study the roles of each aggrecan domain in product biosynthesis, GAG chain modification and product secretion, we generated a number of recombinant constructs carrying different domains of aggrecan: G1, G2, G3, IGD, KS, CS, G1–G2, G1–G3, G1–CS, G1–KS, IGD–G2, G1CSD, miniaggrecan, CS–G3 and G3–KS. The structures of these constructs are shown in detail in Figure 1. To allow direct comparisons of expression and secretion, each construct contained the same leading peptide for product targeting. This leading peptide, originally obtained from link protein, contains an epitope recognized by the monoclonal antibody 4B6 [21]. The mammalian expression vectors pcDNA1, pcDNA3 and pCR3.1 are driven by a CMV (cytomegalovirus) promoter. These constructs were used in transfection of COS-7 cells using Lipofectin, as described in the Experimental section. The growth medium and cell lysate were harvested and analysed on Western blots probed with antibody 4B6 and visualized with a ECL kit (Amersham) according to the manufacturer's instructions.

G3 enhances GAG modification of KS and CS domains, and facilitates product secretion

It has been reported in studies using CHO cells that an aggrecan G3 construct was well synthesized and secreted into the culture

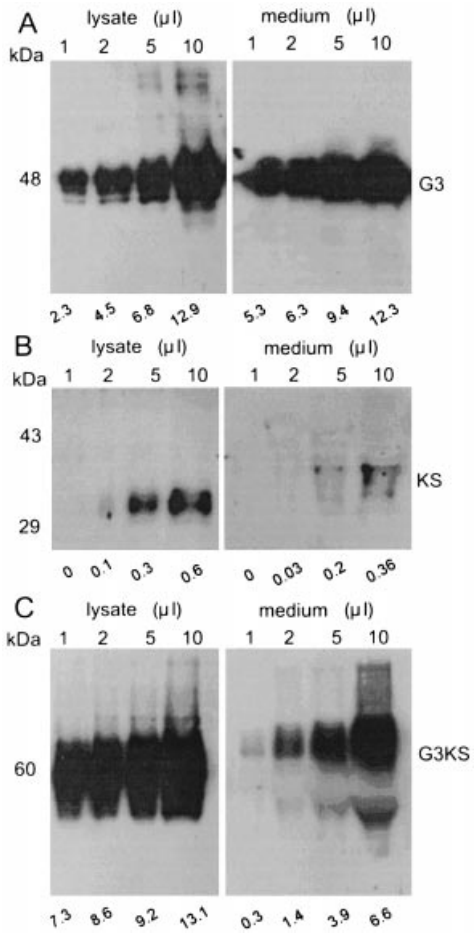

Figure 2 G3 enhances KS domain synthesis and secretion

COS-7 cells were transiently transfected with the recombinant constructs G3 (**A**), KS (**B**) and G3-KS (**C**). The cultures were maintained in DMEM complemented with 5% (v/v) FBS at 37 °C in an incubator for 3 days. Cell lysate and culture media were subjected to SDS/PAGE (1, 2, 5 and 10 μl per well) in a 10% (w/v) polyacrylamide gel. The separated proteins were then transblotted on to a nitrocellulose membrane and probed with the monoclonal antibody 4B6. Products of G3 ($\sim$ 48 kDa) were synthesized and secreted into the culture medium. Products of KS ($\sim$ 30–35 kDa) were weakly synthesized, secreted into the culture medium and modified by GAGs, resulting in smear bands in the gel. Products of G3–KS ($\sim$ 50–70 kDa) were heavily modified by GAGs, resulting in strong smear bands in the gel. The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$). Due to a limiting sensitivity and saturation effect of the film, this reading does not correlate proportionally to increases in volume loaded on to each well.

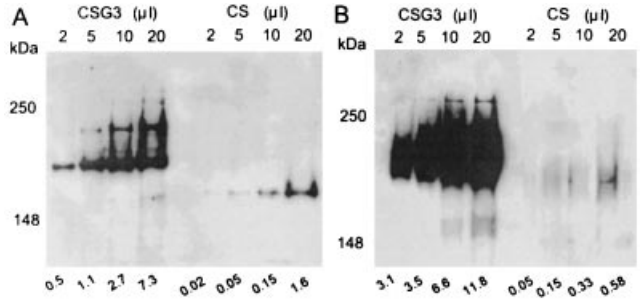

Figure 3 G3 enhances CS domain synthesis and secretion

COS-7 cells were transiently transfected with the recombinant constructs CS and CSG3, as described in the legend to Figure 2. Cell lysate (**A**) and culture media (**B**) were analysed on Western blots (2, 5, 10 and 20 μl per well) probed with antibody 4B6. Products of CS–G3 ($\sim$ 180 kDa) were well synthesized and secreted into the culture medium, while the products of CS ($\sim$ 90 kDa) were poorly synthesized and weakly secreted. The products of CS–G3 were heavily modified by GAG chains, resulting in strong smear bands on the blot ($\sim$ 180–200 kDa), while the products of CS were weakly modified by GAG chains. The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$).

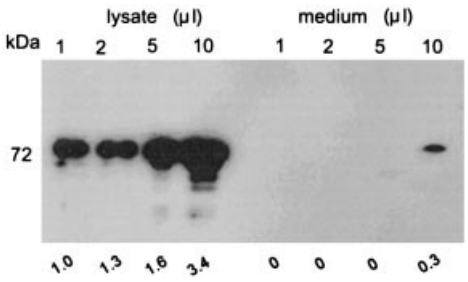

Figure 4 G1 inhibits G3 secretion

Cells were transiently transfected with the recombinant construct G1–G3. Cell lysate and culture medium were analysed on Western blots (1, 2, 5 and 10 μl per well) probed with antibody 4B6. The products ($\sim$ 72 kDa) were well synthesized, but poorly secreted into the culture medium. The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$).

medium, while a G1 construct was poorly synthesized and weakly secreted into the culture medium [22]. Therefore, in our study of the role of this and other aggrecan domains in product biosynthesis, modification by GAG chains and secretion, we first confirmed that an aggrecan G3 construct was well synthesized and secreted into the culture medium when expressed in COS-7 cells (Figure 2A). The KS construct was expressed in the same way; it was weakly synthesized and secreted into the culture medium, and weakly modified by GAG chains (Figure 2B). When the KS domain was linked to the G3 construct, the resulting G3–KS construct was well synthesized and strongly modified by GAG chains. A high proportion of products appeared in the culture medium (Figure 2C).

We further tested the role of the G3 domain in GAG modification of the CS domain sequence. A construct containing the G3 domain and a fragment of the CS domain (nucleotides 3838–5580, equivalent to 51% of the entire CS domain of aggrecan [12,13]) was expressed in COS-7 cells. The CS construct containing the leading peptide and the CS fragment was used as a control. The experiment indicated weak synthesis and secretion of the CS construct (Figure 3A), while the CS–G3 construct was well synthesized and secreted (Figure 3B). When a G1–G3 construct was expressed in the same way, the construct produced a noticeable amount of protein in the cell lysate (Figure 4). Surprisingly, little was observed in the medium, suggesting that the G1 domain exerts a strong inhibitory effect on secretion.

The decrease in GAG modification and inhibition of product secretion caused by G1 is partially abolished by the CS domain, and enhanced by G3

To test further the hypothesis concerning the effect of G1 on secretion, we performed a product secretion assay using a G1 construct. The assay revealed that the G1 products were well

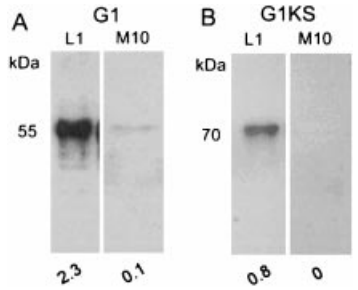

Figure 5 G1 inhibits KS domain secretion

Cells were transiently transfected with the G1 and G1–KS constructs. Cell lysate (L1 ; 1 μl) and culture medium (M10; 10 μl) were analysed on a Western blot probed with antibody 4B6. Products of the G1 construct (55 kDa) were synthesized, but not secreted into the culture medium (**A**). Products of G1–KS were well produced (detected in the cell lysate) but poorly modified by GAG chains, and could not be detected in the culture medium (**B**). The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$).

synthesized, but that the product was barely detected in the culture medium (Figure 5A). We then tested the effect of the G1 domain on GAG modification of the KS sequence. COS-7 cells were transiently transfected with the construct G1–KS. Cell lysate and culture medium were harvested and analysed on a Western blot probed with antibody 4B6. The presence of G1 inhibited the addition of GAG chains to the G1–KS core protein, and completely prevented secretion of the G1–KS product (Figure 5B).

Further, we tested whether G1 could inhibit GAG modification of the CS domain sequence and inhibit product secretion. To do this, we expressed a G1–CS construct in COS-7 cells. The G1 domain only partially inhibited GAG modification of the CS domain sequence and partially decreased product secretion as compared with the control construct CS (Figure 6A). In other words, the inhibitory effect of G1 on product secretion was partially overcome by the presence of the CS domain sequence. The G1 domain was then linked to a small fragment of the CS domain sequence (CSD) in the presence or the absence of the G3 domain. The resulting constructs, G1–CSD and G1–CSD–G3, were expressed. The core protein of G1–CSD was not modified by GAG chains, and the product was not secreted (Figure 6B), while the core protein of G1–CSD–G3 was modified by GAG chains and the products were well secreted (Figure 6C). Treatment of the mini-aggrecan product with chondroitinase ABC resulted in a core protein with a molecular mass of 130 kDa.

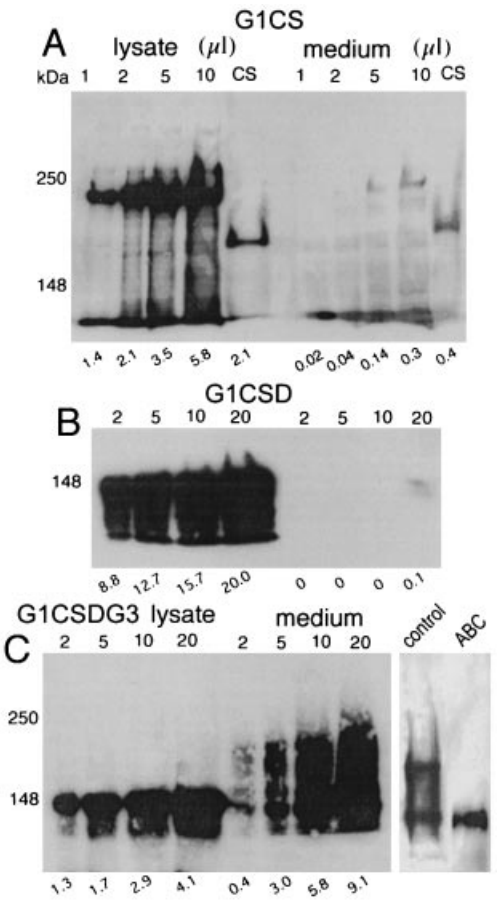

Figure 6 Inhibition of secretion by G1 is overcome by the CS domain and enhanced by G3

Cells were transiently transfected with constructs G1–CS (**A**), G1–CSD (**B**) (where CSD is a small fragment of the CS domain sequence) and mini-aggrecan (**C**) (G1–CSD–G3). Cell lysate and culture medium were analysed on Western blots probed with antibody 4B6. For comparison, the lysate (20 μl) and culture medium (20 μl) of cells containing products of construct CS were loaded as controls in (**A**). The products of G1–CS ($\sim$ 180–200 kDa) were well synthesized, weakly modified by GAG chains and weakly secreted into the culture medium as compared with the products of CS. The products of G1–CSD were well synthesized, but barely detected in the culture medium. The products of mini-aggrecan (G1–CSD–G3) were well synthesized, modified by GAG chains and secreted into the medium. The product of mini-aggrecan in the culture medium was also purified using 4B6 and treated with chondroitinase ABC, and the product with (ABC) or without (control) chondroitinase ABC treatment was analysed by Western blotting. The core protein of the mini-aggrecan was detected with a molecular mass of 130 kDa. The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$).

G2 inhibits product secretion

Since G1 has an inhibitory effect on product secretion, and since G2 exhibits extensive homology with the tandem repeats of G1, we reasoned that G2 might also inhibit product secretion. The G2 construct was expressed in COS-7 cells, and cell lysate and culture medium were analysed on Western blots probed with antibody 4B6. The product was poorly synthesized and was not observed in the culture medium (Figure 7A). A time course study of G2 expression (Figure 7B), with the controls of G1 (Figure 7C) and G3 (Figure 7D), indicated that the products of G1 and G2 accumulated in the cell lysate and reached equilibrium after 2 days of transfection, while the product of G3 reached equilibrium 1 day later. Notably, when the G2 domain was linked to the G1 construct, producing G1–G2, G2 did not inhibit the biosynthesis of G1. As a result, the G1–G2 products were well produced, but, as expected, they were not secreted into the culture medium (Figure 7E).

Finally, we investigated the role of the IGD in product biosynthesis and secretion. COS-7 cells were transiently transfected with the IGD construct. Analysis of the cell lysate and culture medium indicated not only that the IGD products were well produced and secreted into the culture medium, but also that the core proteins were heavily modified by GAG chains (Figure 7F). A search of the sequence of the IGD revealed a unique pair of Ser-Gly residues, a potential sugar attachment site, coded by nucleotides 1055–1060. This suggests that this site is modified by GAG chain attachment. In light of the proximity of IGD to G2 (G2 is directly C-terminal to IGD), we chose to investigate the potential effect of the G2 domain on IGD biosynthesis, modification by GAG chains and product secretion.

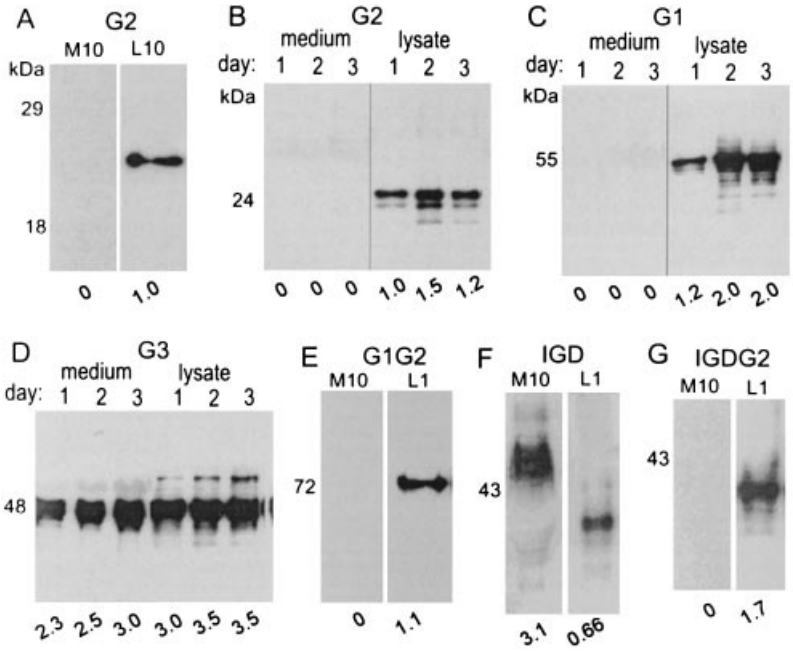

Figure 7 Effects of the G2 domain on GAG chain attachment and product secretion

Cells were transiently transfected with the constructs G2, G1, G3, G1–G2, IGD and IGD–G2, as described in the legend to Figure 2. Cell lysate (L1 and L10; 1 and 10 μl respectively) and culture medium (M10; 10 μl) were analysed on Western blots and probed with antibody 4B6. Products of G2 ($\sim$ 22 kDa) were synthesized, but were not secreted into the culture medium (**A**). The time course of the expression of G2 was also analysed (**B**), while the expression of G1 (**C**) and G3 (**D**) served as controls for the time course study (10 μl samples). Products of G1–G2 were well synthesized, but were not secreted into the culture medium (**E**). Products of IGD were synthesized and modified by GAG chains, resulting in smear bands in the gel, and secreted into the culture medium. The sizes of protein in the lysate were approx. 40–45 kDa, while those in the culture medium were approx. 45–65 kDa (**F**). The products of IGD–G2 were well synthesized and modified by GAG chains, but they were not secreted (**G**). The relative densities of each band after Western blot development are shown below each gel ($\times 10^3$).

An IGD–G2 construct was generated and expressed as described above. Western blot analysis indicated that the G2 domain completely prevented the secretion of IGD (Figure 7G). However, G2 did not inhibit biosynthesis, and seemed to reduce GAG chain modification of the IGD–G2 core protein only.

DISCUSSION

Our work reported here represents the first comprehensive study of the role of each aggrecan domain in product processing. We have demonstrated that the G1 and G2 domains reduce modification by GAG chains of the core proteins and inhibit product secretion. There are also two components that stimulate product secretion: the G3 domain and the GAG-chain-modified sequences (CS domain, KS domain and IGD). The G3 domain not only stimulates product secretion, but also promotes the attachment of GAG chains. The KS and CS domain sequences are poorly modified by GAG chains in the absence of the G3 domain. The G3 domain, however, is not able to stimulate the secretion of G1 sufficiently. For instance, the products of G1–G3 were barely detected in the culture medium. Only in the presence of a fragment of the CS domain sequence (e.g. the CSD fragment in the G1–CSD–G3 construct), where the G3 domain facilitates GAG chain attachment to the CS fragment, resulting in the presence of two positive components for product secretion, can the inhibitory effects of G1 be abolished. The fact that the endogenous aggrecan, which contains two components (G1 and G2 domains) for the inhibition of product secretion and two components (G3 domain and modified GAG chains) for product secretion, is secreted into the ECM implies that the G3 and CS domains, by working together, play a predominant role in product secretion. In our experience, we have never observed secretion of the unmodified core proteins of mini-aggrecan, mini-versican, aggrecan G1–CS construct or versican G1–CS construct. These results indicate that GAG chain attachment is imperative for secretion of these proteoglycans. Further support for this is provided by our observation that the inhibition of product secretion by G1 was overcome by a larger CS domain sequence (G1–CS), but not by a smaller CS domain sequence (G1–CSD). It has been reported that the enhancement by the G3 domain of product secretion occurs through its interaction with HSP25 (heat-shock protein of 25 kDa) [22]. Each protein module also enhances product secretion [23,24]. The mechanism of our observation that modification by GAG chains stimulates product secretion awaits future investigation.

An interesting question then arises: why are chondrocytes programmed to produce two components for inhibition of product secretion? Our findings that G3, in the absence of a CS domain fragment, was not able to stimulate G1 secretion in the construct G1–G3, and that the G1–CS core protein was weakly modified by GAG chains and the GAG-modified products were secreted, seem to indicate that the cells are programmed to produce 'perfect' aggrecan molecules. Aggrecan core protein cannot be mistakenly secreted into the matrix following mRNA translation, as the G1 and G2 domains will inhibit its secretion. Only after the aggrecan core protein is correctly modified by GAG chains could the proteoglycan be secreted. This programming is very important for cartilage development and maintenance, since the GAG chains play important roles in

retaining water and negative charges and in producing a load-bearing cartilage.

The presence of the G2 domain may be essential, since cells have to completely block secretion of any imperfect aggrecan. Indeed, without the G2 domain, the G1–CS core protein was weakly modified and the product was weakly secreted. One may imagine that the products of such under-modification by GAG chains would be harmful in cartilage development and maintenance. This hypothesis is supported not only by our experimental results, but also by previous studies on other members of this CS proteoglycan family, which indicated that these proteoglycans are essential for tissue development and growth [25–28]. Aggrecan contains a G2 domain, and its core protein is heavily modified by GAG chains, while other family members (i.e. versican, neurocan and brevican) do not contain a G2 domain and their core proteins are not heavily modified by GAG chains. This should not lead to an incorrect assumption that G2 enhances modification by GAG chains. It is the G3 domain, and not the G2 domain, that promotes GAG chain attachment. The G2 domain not only inhibits product secretion, but also reduces modification by GAG chains. Its effect on product secretion seems to be predominant, in order to inhibit secretion of insufficiently modified aggrecan. As a result, only those molecules sufficiently modified by GAG chains are secreted into the ECM. The G3 domain, on the other hand, enhances GAG chain attachment and stimulates product secretion. The importance of the G3 domain is seen in nanomelia, a fatal genetic disease of the chicken. Nanomelic aggrecan contains a premature stop codon at the 3′ end of the CS domain sequence (5′ to the G3 domain), and the mutant aggrecan core protein is neither modified by GAG chains nor secreted into the ECM [29,30]. Another fatal genetic disease that demonstrates the importance of the CS and G3 domains is the cartilage deficiency in mice whose aggrecan contains a 7 bp deletion in exon 5 (tandem repeat 1 motif of the G1 domain). As a result, the truncated aggrecan can no longer be secreted into the matrix [26,31].

It has long been known that the G2 domain is structurally similar to the tandem repeats of the G1 domain. While the G1 domain binds to link protein via the IgG-like motif and binds hyaluronan via the tandem repeats, forming ternary complexes to maintain a stable matrix network in cartilage, the G2 domain does not bind to hyaluronan [32] and, so far, no molecule has been found to interact with G2. The function of the G2 domain remains unknown. Our report on the role of the G2 domain in aggrecan processing has, for the first time, shed some light on the biology of this domain.

Another novel finding in our studies is the modification of the IGD by GAG chains. There is only one potential site on the IGD for GAG chain attachment. Our observation that the IGD construct was synthesized and modified by GAG chains, and that the product was secreted into the culture medium, implies that the potential site is actually modified by GAG chains in the COS-7 cell line. This could only be achieved by using a small recombinant construct. In a large construct, which may contain many sites for GAG chain attachment, one cannot distinguish the presence or absence of one particular GAG chain. It is not yet known whether this site is modified by GAG chains in endogenous aggrecan.

Thus, given the diversity of effects exerted by aggrecan domains on product processing, it is likely that each domain is involved in the processing and secretion of the mature aggrecan product. In the model we propose, the G1 and G2 domains reduce modification by GAG chains, which is overcome by the G3 domain. G3, through promotion of GAG chain attachment, overcomes the inhibitory effects of G1 and G2 on product secretion. The attachment of GAG chains to the IGD, the KS domain and the CS domain, which allows aggrecan to generate stable matrix meshwork, is also a prerequisite for aggrecan secretion.

We thank Dr Paul F. Goetinck for the monoclonal antibody 4B6, the aggrecan CS domain fragment and the aggrecan G3 domain, and Dr Yeqi Yao for technical help. This work was supported by grant MOP-13730 from the Medical Research Council of Canada (to B.B.Y.). M.E.A. is on sabbatical from the University of Calgary. B.B.Y. is a Scholar of the Arthritis Society of Canada. C.K. is supported by the Canadian Arthritis Network Graduate Scholarships in Science and Technology.

REFERENCES

1 Wight, T. N., Heinegård, D. K. and Hascall, V. C. (1991) Proteoglycans: structure and function. In Cell Biology of Extracellular Matrix (Hay, E. D., ed.), pp. 45–78, Plenum Press, New York
2 Watanabe, H., Gao, L., Sugiyama, S., Doege, K., Kimata, K. and Yamada, Y. (1995) Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene structure and promoter sequence. Biochem. J. **308**, 433–440
3 Goetinck, P. F. (1991) Proteoglycans in development. Curr. Top. Dev. Biol. **25**, 111–131
4 Toole, B. P. (1990) Hyaluronan and its binding proteins, the hyaladherins. Curr. Opin. Cell Biol. **2**, 839–844
5 Vilim, V. and Fosang, A. J. (1994) Proteoglycans isolated from dissociative extracts of differently aged human articular cartilage: characterization of naturally occurring hyaluronan-binding fragments of aggrecan. Biochem. J. **304**, 887–894
6 Lohmander, L. S., Dahlberg, L., Ryd, L. and Heinegard, D. (1989) Increased levels of proteoglycan fragments in knee joint fluid after injury. Arthritis Rheum. **32**, 1434–1442
7 Lohmander, L. S., Hoerner, L. A., Dahlberg, L., Roos, H., Bjornsson, S. and Lark, M. W. (1993) Stromelysin, tissue inhibitor and proteoglycan fragments in human knee joint fluid after injury. J. Rheumatol. **20**, 1362–1368
8 Deak, F., Kiss, I., Sparks, K. J., Argraves, W. S., Hampikian, G. and Goetinck, P. F. (1986) Complete amino acid sequence of chicken cartilage link protein deduced from cDNA clones. Proc. Natl. Acad. Sci. U.S.A. **83**, 3766–3770
9 Margolis, R. U. and Margolis, R. K. (1994) Aggrecan-versican-neurocan family proteoglycans. Methods Enzymol. **245**, 105–126
10 Sai, S., Tanaka, T., Kosher, R. A. and Tanzer, M. L. (1986) Cloning and sequence analysis of a partial cDNA for chicken cartilage proteoglycan core protein. Proc. Natl. Acad. Sci. U.S.A. **83**, 5081–5085
11 Zheng, J., Luo, W. and Tanzer, M. L. (1998) Aggrecan synthesis and secretion. A paradigm for molecular and cellular coordination of multiglobular protein folding and intracellular trafficking. J. Biol. Chem. **273**, 12999–13006
12 Chandrasekaran, L. and Tanzer, M. L. (1992) Molecular cloning of chicken aggrecan. Structural analyses. Biochem. J. **288**, 903–910
13 Li, H., Schwartz, N. B. and Vertel, B. M. (1993) cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia. J. Biol. Chem. **268**, 23504–23511
14 Yang, B. B., Zhang, Y., Cao, L. and Yang, B. L. (1997) Aggrecan and link protein affect cell adhesion to culture plates and to type II collagen. Matrix Biol. **16**, 541–561
15 Cao, L., Zhang, Y. and Yang, B. B. (1998) Expression of the G1 domain of aggrecan interferes with chondrocyte attachment. Matrix Biol. **17**, 379–392
16 Cao, L. and Yang, B. B. (1999) Chondrocyte apoptosis induced by aggrecan G1 domain as a result of decreased cell adhesion. Exp. Cell Res. **246**, 527–537
17 Stirpe, N. S., Argraves, W. S. and Goetinck, P. F. (1987) Chondrocytes from the cartilage proteoglycan-deficient mutant, nanomelia, synthesize greatly reduced levels of the proteoglycan core protein transcript. Dev. Biol. **124**, 77–81
18 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19 Zhang, Y., Cao, L., Yang, B. L. and Yang, B. B. (1998) The G3 domain of versican enhances cell proliferation via epidermal growth factor-like motifs. J. Biol. Chem. **273**, 21342–21352
20 Zhang, Y., Cao, L., Yang, B. L. and Yang, B. B. (1998) The G3 domain of versican inhibits mesenchymal chondrogenesis via the EGF-like motifs. J. Biol. Chem. **273**, 33054–33063
21 Binette, F., Cravens, J., Kahoussi, B., Haudenschild, D. R. and Goetinck, P. F. (1994) Link protein is ubiquitously expressed in non-cartilaginous tissues where it enhances and stabilizes the interaction of proteoglycans with hyaluronic acid. J. Biol. Chem. **269**, 19116–19122

22 Luo, W., Kuwada, T. S., Chandrasekaran, L., Zheng, J. and Tanzer, M. L. (1996) Divergent secretory behavior of the opposite ends of aggrecan. J. Biol. Chem. **271**, 16447–16450
23 Day, J. M., Murdoch, A. D. and Hardingham, T. E. (1999) The folded protein modules of the C-terminal G3 domain of aggrecan can each facilitate the translocation and secretion of the extended chondroitin sulfate attachment sequence. J. Biol. Chem. **274**, 38107–38111
24 Domowicz, M. S., Pirok, III, E. W., Novak, T. E. and Schwartz, N. B. (2000) Role of the C-terminal G3 domain in sorting and secretion of aggrecan core protein and ubiquitin-mediated degradation of accumulated mutant precursors. J. Biol. Chem. **275**, 35098–35105
25 Vertel, B. M., Walters, L. M., Grier, B., Maine, N. and Goetinck, P. F. (1993) Nanomelic chondrocytes synthesize, but fail to translocate, a truncated aggrecan precursor. J. Cell Sci. **104**, 939–948
26 Watanabe, H., Kimata, K., Line, S., Strong, D., Gao, L-Y., Kozak, C. A. and Yamada, Y. (1994) Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the aggrecan gene. Nat. Genet. **7**, 154–157
27 Kerr, R. S. E. and Newgreen, D. F. (1997) Isolation and characterization of chondroitin sulfate proteoglycans from embryonic quail that influence neural crest cell behavior. Dev. Biol. **192**, 108–124
28 Mjaatvedt, C. H., Yamamura, H., Capehart, A. A., Turner, D. and Markwald, R. R. (1998) The Cspg2 gene, disrupted in the hdf mutant, is required for right cardiac chamber and endocardial cushion formation. Dev. Biol. **202**, 56–66
29 Primorac, D., Stover, M. L., Clark, S. H. and Rowe, D. W. (1994) Molecular basis of nanomelia, a heritable chondrodystrophy of chicken. Matrix Biol. **14**, 297–305
30 Vertel, B. M., Grier, B. L., Li, H. and Schwartz, N. B. (1994) The chondrodystrophy, nanomelia: biosynthesis and processing of the defective aggrecan precursor. Biochem. J. **301**, 211–216
31 Wai, A. W., Ng, L. J., Watanabe, H., Yamada, Y., Tam, P. P. and Cheah, K. S. (1998) Disrupted expression of matrix genes in the growth plate of the mouse cartilage matrix deficiency (cmd) mutant. Dev. Genet. **22**, 349–358
32 Watanabe, H., Cheung, S. C., Itano, N., Kimata, K. and Yamada, Y. (1997) Identification of hyaluronan-binding domains of aggrecan. J. Biol. Chem. **272**, 28057–28065

Received 18 August 2000/1 November 2000; accepted 1 December 2000