THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 10, Issue of March 9, pp. 7507–7517, 2001

Printed in U.S.A.

Mechanisms Underlying Preferential Assembly of Heparan Sulfate on Glypican-1*

Received for publication, September 11, 2000, and in revised form, November 29, 2000
Published, JBC Papers in Press, December 5, 2000, DOI 10.1074/jbc.M008283200

Robert L. Chen‡ and Arthur D. Lander§

From the Department of Developmental and Cell Biology and Developmental Biology Center, University of California, Irvine, California 92697

Glypicans are major cell surface heparan sulfate proteoglycans, the structures of which are characterized by the presence of a cysteine-rich globular domain, a short glycosaminoglycan (GAG) attachment region, and a glycosylphosphatidylinositol membrane anchor. Despite strong evolutionary conservation of the globular domains of glypicans, no function has yet been attributed to them. By using a novel quantitative approach for assessing proteoglycan glycosylation, we show here that removal of the globular domain from rat glypican-1 converts the proteoglycan from one that bears ~90% heparan sulfate (HS) to one that bears ~90% chondroitin sulfate. Mutational analysis shows that sequences at least 70 amino acids away from the glypican-1 GAG attachment site are required for preferential HS assembly, although more nearby sequences also play a role. The effects of the glypican-1 globular domain on HS assembly could also be demonstrated by fusing this domain to sequences representing the GAG attachment sites of other proteoglycans or, surprisingly, simply by expressing the isolated globular domain in cells and analyzing effects either on an exogenously expressed glypican-1 GAG attachment domain or on endogenous proteoglycans. Quantitative analysis of the effect of the globular domain on GAG addition to proteoglycan core proteins suggested that preferential HS assembly is achieved, at least in part, through the inhibition of chondroitin sulfate assembly. These data identify the glypican-1 globular domain as a structural motif that potently influences GAG class determination and suggest that an important role of glypican globular domains is to ensure a high level of HS substitution of these proteoglycans.

Proteoglycans (PGs)[1] influence a variety of cellular and physiological activities including cell proliferation, cell adhesion, blood coagulation, and wound repair (1). In most cases, the biological activities of PGs depend on the molecular interactions of their covalently attached glycosaminoglycan (GAG) chains. The major GAGs that are found on PGs, heparan sulfate (HS) and chondroitin sulfate (CS),[2] are both synthesized via similar routes, involving the stepwise addition of monosaccharides to the side chain of a Ser residue N-terminal to a Gly residue. HS and CS both begin with the same tetrasaccharide moiety, GlcUAβ1–3Galβ1–3Galβ1–4Xylβ1-Ser. To the terminal GlcUA is attached either α-GlcNAc or β-GalNAc, depending on whether HS or CS, respectively, is to be synthesized. Subsequently, copolymerases elongate HS or CS backbones through the repetitive addition of GlcUA and the appropriate *N*-acetylhexosamine, and other modification reactions, sulfation and epimerization, add complexity and functionality to the chains (1–3).

As HS and CS have divergent structures and distinct biochemical properties, it is important that cells appropriately regulate which GAGs they attach to different PG core proteins. Based on the biosynthetic scheme just described, whether a PG carries HS or CS should be governed by either the accessibility or activity of the two enzymes, termed α-*N*-acetylglucosaminyltransferase I and β-*N*-acetylgalactosaminyltransferase I (β-GalNAc-TI), that are capable of adding the fifth sugar onto the common GlcUAβ1–3Galβ1–3Galβ1–4Xylβ1 "linkage" tetrasaccharide (4, 5). Clearly, much of the information that regulates this step must be encoded in the PG core protein, since cells that are capable of making both types of GAGs reliably place the correct ones onto appropriate core proteins. For example, cores that bear HS include glypicans, syndecans, perlecan, and agrin; cores that bear CS include members of the aggrecan and decorin families, among others. Some PGs, such as syndecan-1, commonly possess both HS and CS at different sites on the protein (6, 7), whereas others, such as serglycin, bear either CS or HS, depending upon the cell type in which they are expressed (8, 9). Finally, some core proteins may be produced as a mixture of molecules possessing and lacking GAGs, or may lack GAGs entirely in some cell types. Recent data suggest that such "naked" core proteins actually possess the GlcUAβ1–3Galβ1–3Galβ1–4Xylβ1 linkage tetrasaccharide at their GAG attachment sites (10). Accordingly, they may be referred to as glycosylated but not glycanated (*i.e.* not GAG-extended).

Attempts to elucidate signal elements in PG core proteins that regulate glycanation have focused attention on amino acid sequences close to GAG attachment sites. Zhang and co-workers (11, 12) fused short segments (typically <30 amino acids) of the core proteins of various PGs to the IgG-binding domain of protein A and expressed these constructs, and various mutant

* This work was supported by National Institutes of Health Grant NS26862 (to A. D. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Member of the Medical Scientist Training Program at the University of California, Irvine.

§ To whom correspondence and reprint requests should be addressed: Dept. of Developmental and Cell Biology, 5207 Biological Sciences II, University of California, Irvine, CA 92697–2300. Tel.: 949-824-1008; Fax: 949-824-1083; E-mail: adlander@uci.edu.

[1] The abbreviations used are: PG(s), proteoglycan(s); β-GalNAc-TI, β-*N*-acetylgalactosaminyltransferase I; AP, alkaline phosphatase; CHO, Chinese hamster ovary; CS, chondroitin sulfate; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; GPI, glycosylphosphatidylinositol; HS, heparan sulfate; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; NBT, nitro blue tetrazolium; BCIP, 5-bromo-4-chloro-3-indolyl phosphate.

[2] In this report we treat dermatan sulfate as a subset of CSs, since the former is simply a highly epimerized variant of the latter, and epimerization takes place at a later biosynthetic stage than the ones considered here.

and chimeric forms of them, in mammalian cells. The secreted fusion proteins were then isolated and analyzed for their proportion of HS *versus* CS. From these studies it was concluded that particular sequence elements act as important enhancers of HS biosynthesis, such that in the absence of such elements glycanation with CS greatly predominated. The elements that favored HS formation included having two or more Ser-Gly dipeptides in close proximity to each other, having a nearby (within 10 amino acids) cluster of acidic amino acids, and having a tryptophan residue immediately following a Ser-Gly GAG attachment site. Some of the same HS-enhancing elements have also been reported by others (*e.g.* see Ref. 13).

Although it is likely that the motifs elucidated by such studies play important roles in enhancing HS assembly on intact PGs, it is noteworthy that such elements are not sufficient to drive efficient glycanation with HS. In the studies of Zhang and co-workers (11, 12), for example, even the fusion proteins with the most favorable sequences rarely carried more than 60% of their GAG as HS. In contrast, several types of PGs are often glycanated exclusively with HS *in vivo*. A particularly striking example is the cell surface PG glypican-1. It appears to carry purely HS *in vivo* (14–16), yet a protein A fusion containing a 24-amino acid fragment derived from the glypican-1 GAG attachment region carried only 20% HS (12). This result was obtained despite the fact that this 24-amino acid region contains three consecutive Ser-Gly dipeptides and two nearby clusters of acidic amino acids. Such findings imply that PG core proteins must contain, outside of their immediate GAG attachment sites, information that regulates glycanation. In the case of glypican-1, such information must be particularly important for determining whether HS or CS is produced.

Glypican core proteins, of which six have been identified in vertebrates, two in *Drosophila*, and one in *Caenorhabditis elegans* (17), share a unique, highly conserved structure, in which a large N-terminal globular domain containing 14 cysteine residues is followed by a GAG attachment domain that includes multiple Ser-Gly dipeptide sites and finally a glycosylphosphatidylinositol (GPI)-anchor that tethers the molecule to the plasma membrane. To date, no binding properties or biochemical activities have been attributed to the N-terminal globular domain of glypicans, despite the fact that these domains display a high degree of sequence conservation both among animal species and between glypican family members. In the present study we demonstrate that the N-terminal globular domain of glypican-1 is a potent enhancer of preferential HS glycanation, capable of driving the glypican-1 GAG attachment domain to assemble HS almost exclusively. We further show that the glypican-1 globular domain can shift the balance of HS glycanation on other GAG attachment domains to which it is experimentally fused and, surprisingly, even on PGs to which it is not fused but is simply coexpressed. Finally, we provide data suggesting that the mechanism by which the globular domain promotes HS assembly involves inhibition of the initiation of CS chains.

MATERIALS AND METHODS

Cell Culture—COS-7 and Chinese hamster ovary (CHO-K1) cells were obtained from the American Type Culture Collection (CCL-61, Manassas, VA). COS cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal bovine serum (HyClone, Salt Lake City, UT), and CHO cells were grown in Ham's F-12 medium supplemented with 7.5% (v/v) fetal bovine serum. Both cell types were grown with penicillin G (100 units/ml) and streptomycin sulfate (100 μg/ml) at 37 °C in a 5% CO_2 atmosphere. Cells were passaged every 3–4 days.

Fusion Protein Construction—The vector *APtag-6His* was produced by transferring the *Hin*dIII to *Xba*I fragment from the *APtag-2* vector (18) containing the AP coding sequence followed by the *APtag-2* polylinker sequence into pcDNA3 (Invitrogen, Carlsbad, CA). Next, a segment encoding the peptide sequence RGSHHHHHH, followed by a termination codon, was produced by annealing the oligonucleotides 5′-ATCGTGGTTCTCATCACCATCACCATCACTAACTCGAG-3′ and 5′CTCGAGTTAGTGATGGTGATGGTGATGAGAACCACGAT-3′. This fragment was inserted into the unique *Hpa*I site at the end of the AP coding sequence. This insertion interrupts the native AP termination codon and thus allows translation through the polyhistidine tag. This construct and all subsequent constructs were sequenced to confirm their identity.

The plasmid encoding GPC(N1/N2/G)-AP was produced by isolating nucleotides 9–1459 of rat glypican-1 cDNA (see Ref. 19; GenBank™ accession number L34067) with *Bam*HI and *Nar*I. The segment containing nucleotides 1459–1641 was amplified by PCR using the primers 5′-^{1449}CGTTTACGTGGCGCC-3′ and 5′-^{1641}CGGGATCCGGTCTTCT-GTCCCTC-3′. The PCR product was digested with *Nar*I and *Bam*HI. The released glypican-1 fragment and the PCR fragment were subsequently inserted, in the proper orientation, into the *Bgl*II site of *APtag-6His*.

The plasmid encoding GPC(N1/N2)-AP was generated with two PCR products amplified from *APtag-6His*/GPC(N1/N2/G)-AP. The first PCR product was generated with 5′-^{1300}TGGAATGGGATTTCCAA-3′ and 5′-ATCGCCACCGTAGGCGCCACG-3′(the fourth nucleotide of this sequence corresponds to position 1473 of native glypican-1) and digested with *Kpn*I. The second PCR product was generated from the AP coding sequence with 5′-ATCAGATCTATCATCCCAGTTGAGGAG-3′ and 5′-GGAAAATCCTAGGACCGT-3′ and digested with *Bam*HI. Both fragments, when joined at their blunt ends, generate an *Eco*RV site at the junction between glypican-1 and AP. The digested PCR products were ligated with *APtag-6His*/GPC(N1/N2/G)-AP that had previously been cut with *Kpn*I and *Bam*HI.

The plasmid encoding GPC(G)-AP was constructed by replacing the *Not*I to *Nar*I segment from the glypican-1 coding region of *APtag-6His*/GPC(N1/N2/G)-AP with an adapter sequence, generated by annealing 5′-GGCCGCCGCGCTAGTCGCCTGCGCCCGCGGGGACCCCGCCAG-CAAGGG-3′ and 5′-CGCCCTTGCTGGCGGGGTCCCCGCGGGGCGC-AGGCGACTAGCGCGGC-3′. This adapter sequence replaces a small part of the signal peptide that was excised in the *Not*I to *Nar*I digest and connects the signal peptide directly to region G.

The plasmid encoding AP-GPC(G) was constructed by amplifying the glypican-1 cDNA with 5′-CGGGATCCGGCGCCTACGGTGGCAAT-3′ (the ninth nucleotide of this sequence corresponds to position 1459 of native glypican-1) and 5′-^{1641}CGGGATCCGGTCTTCTGTCCCTC-3′, digesting the PCR product with *Bam*HI, and ligating it to the *APtag-4* (18) vector that had been pre-digested with *Bam*HI in the proper orientation.

The plasmid encoding GPC(N2/G)-AP was made by replacing the *Not*I to *Kpn*I segment from the glypican-1 coding region of *APtag-6His*/GPC(N1/N2/G)-AP with an adapter sequence generated by annealing 5′-GGCCGCCGCGCTAGTCGCCTGCGCCCGCGGGGACCCCGCCAG-CAAGTGGAATGGGATTTCCAAGGGCCGGTAC-3′ and 5′-CGGCCCT-TGGAAATCCCATTCCAGCACTTGCTGGCGGGGTCCCCGCGGGGC-GCAGGCGACTAGCGCGGC-3′. This adapter sequence replaces a small part of the signal peptide that was excised in the *Not*I to *Kpn*I digest and connects the signal peptide directly to region N2.

Region N2 mutants N2A, N2B, and N2E were prepared for mutagenesis by subcloning the *Kpn*I to *Nar*I fragment of the glypican-1 cDNA into pBluescript (Stratagene, La Jolla, CA). The segment containing the N2A mutation was generated by replacing the native *Kpn*I to *Bal*I fragment with an adapter sequence consisting of 5′-CTTCCGGAGATC-GAGGGGCGCGGGCTGG-3′ and 5′-CCAGCCCGCGCCCCTCGATCT-CCGGAAGGTAC-3′. This manipulation changes the original amino acid sequence, VMGD, to IEGR. The segment containing the N2B mutation was made by replacing the native *Kpn*I to *Bal*I fragment with an adapter sequence consisting of 5′-CGGTTGG-3′ and 5′-CCAACCG-GTAC-3′. This manipulation replaces the original amino acid sequence, LPEVMGDG, with an Arg residue. The segment containing the N2E mutation was generated first by removing the *Bal*I to *Kas*I fragment, then DNA polymerase I blunting of the overhanging *Kas*I end, followed by blunt-ended ligation. This series of treatments deletes 35 amino acids from region N2. Following mutagenesis, the mutagenized fragments were released from pBluescript with *Kpn*I and *Nar*I then subsequently used to replace the native sequence flanked by *Kpn*I and *Nar*I sites in GPC(N1/N2/G)-AP.

The plasmid encoding N2C was constructed with two PCR products amplified from the glypican-1 cDNA, the first with 5′-^{1300}TGGAATGG-GATTTCCAA-3′ and 5′-CCGAATTCACGTAAACGGTTGGTCA-3′ (the ninth nucleotide of this sequence corresponds to position 1458 of native glypican-1) and the second with 5′-CCGAATTCTACGGTGGCAAT-

GAT-3′ (the ninth nucleotide of this sequence corresponds to position 1465 of native glypican-1) and 5′-^{1641}CGGGATCCGGTCTTCTGTCCCTC-3′. The first and second PCR products were digested with *Eco*RI followed by *Kpn*I and *Bam*HI, respectively. These two PCR fragments, when joined, replace the native *Nar*I site in glypican-1 with an *Eco*RI site. A third fragment containing the remaining N-terminal glypican-1 coding sequence was released with *Hin*dIII and *Kpn*I from *APtag-6His*/GPC(N1/N2/G)-AP. These three DNA fragments were simultaneously ligated into *APtag-6His* vector previously treated with *Hin*dIII and *Bgl*II.

The plasmid encoding N2D was constructed with a PCR product made from the primers 5′-^{1455}ACGTGGCGCCTACGGTGGCATCGAGGGGCGCAATGATGTGGACTTC-3′ and 5′-^{1641}CGGGATCCGGTCTTCTGTCCCTC-3′. This PCR product was intended to replace the native AYGG amino acid sequence with IEGR. However, because of a PCR error the same native sequence was replaced with the tripeptide EGR. This PCR product was digested with *Nar*I and *Bam*HI. Together with a fragment flanked by *Hin*dIII and *Nar*I from *APtag-6His*/GPC(N1/N2/G)-AP, the fragments were ligated into *APtag-6His* vector pretreated with *Hin*dIII and *Bgl*II.

The plasmid encoding secreted AP was generated with an adapter sequence made with 5′-GGCCGCCGCGCTAGTCGCCTGCGCCCGCGGGGACCCCGCCAGCAAGAGCGAT-3′ and 5′-ATCGCTCTTGCTGGCGGGGTCCCCGCGGGCGCAGGCGACTAGCGCGGC-3′. The adapter sequence was ligated with *APtag-6His*/GPC(N1/N2)-AP that had been pretreated with *Not*I and *Eco*RV. The resultant coding sequence contains the glypican-1 N-terminal signal peptide followed by AP, with an *Eco*RV site between the coding regions.

The segment encoding a rat betaglycan GAG attachment domain (amino acids 524–545; GenBank™ accession number M77809) was made with two self-annealing primers 5′-ATCATTGTGGTGCAGGCTCCGTCCCCTGGGGATAGCAGTGGCTGG-3′ and 5′-CCGGATCCCTCGAGGTCTTCATAGCCATCAGGCCAGCCACTGCT-3′. Complementary strands were synthesized using DNA polymerase I. The DNA fragment possesses half of the *Eco*RV site at the 5′ end and a *Bam*HI site at the 3′ end. After synthesis the fragment was digested with *Bam*HI, and the resultant product was ligated into *APtag-6His*/GPC(N1/N2)-AP and *APtag-6His*/AP, both of which had been pretreated with *Eco*RV and *Bgl*II, to produce vectors encoding GPC(N1/N2)-B-AP and B-AP, respectively.

The segment encoding the human decorin GAG attachment domain (amino acids 23–46; GenBank™ accession number M14219) was created with the primers 5′-ATCGGCTTATTTGACTTTATGCTAGAAGATGAGGCTTCTGGG-3′ and 5′-CCGGATCCGAAGTCGCGGTCATCAGGAACTTCTGGGCCTATCCCAGA-3′. Fragment synthesis, digestion, and insertion into plasmids were performed in a manner similar to the betaglycan fragment, producing vectors encoding GPC(N1/N2)-D-AP and D-AP.

Transfection and Harvesting of Fusion Proteins—5×10^5 COS cells or CHO cells were seeded in 100-mm dishes in 5 ml of the appropriate growth medium for 24 h. Prior to transfection, cells were washed in 5 ml of HBSS followed by 5 ml of Opti-MEM (Life Technologies, Inc.). Cells were then transfected with 5 μg of plasmids encoding AP-tagged protein. In experiments in which multiple plasmids were cotransfected into COS cells, 10 μg of AP or GPC(N1/N2)-AP was used and mixed, where indicated, with 1 μg of GPC(G)-AP. Transfection was carried out using LipofectAMINE (Life Technologies, Inc.) in Opti-MEM medium and an incubation time of 12 h. Next, the transfection medium was replaced with 10 ml of the appropriate growth medium, and this medium was allowed to remain in contact with COS cells for 5 days or CHO cells for 3 days. The conditioned medium was then collected, centrifuged, passed through a 0.45-μm filter, supplemented with 2 mM phenylmethylsulfonyl fluoride and 0.05% sodium azide, and stored at 4 °C.

DEAE-Sepharose Chromatography—AP fusion proteoglycans were purified by anion exchange chromatography on DEAE-Sephacel (Amersham Pharmacia Biotech) as described (20), except that detergents and formate- and urea-containing wash buffers were omitted. The AP concentrations of starting and eluted materials were determined from their enzymatic activities (21), and samples were stored at −80 °C with protease inhibitors (2 mM *N*-methylmaleimide, 1 μg/ml pepstatin A, 2 mM phenylmethylsulfonyl fluoride, 1 μg/ml leupeptin), 0.5 mg/ml bovine serum albumin, and 10% glycerol prior to analysis.

GAG Lyase Treatments—Heparitinase was prepared from *Flavobacterium heparinum* by hydroxyapatite chromatography as described by Linker and Hovingh (22). Heparitinase and chondroitinase ABC (100330, Seikagaku America, Falmouth, MA) treatments (1 unit/ml) were carried out in 50 mM Tris phosphate (pH 7.0) containing protease inhibitors for 2 h with 1 pmol/ml DEAE-purified fusion proteoglycans as substrates. Incubation temperatures were 37 °C for chondroitinase and 43 °C for heparitinase. Heparitinase/chondroitinase combination treatments were carried out at 37 °C.

Polyacrylamide Gel Electrophoresis and Chromogenic Detection of Fusion Proteins—Samples were mixed 4:1 with 5-fold concentrated SDS-PAGE sample buffer and subjected (without heating) to electrophoresis in 6% Laemmli gels for 200 V-h. $MgCl_2$ (1 mM) was added to all buffers to stabilize AP activity. Gels were subsequently washed for 5 min in 50 mM Tris-HCl (pH 8.0) containing 1% Triton X-100 and 1 mM $MgCl_2$ and for 5 min in 100 mM diethanolamine (pH 9.8) containing 1 mM $MgCl_2$. Gel staining was carried out in 20 ml of Sigma FAST BCIP/NBT (Sigma B-5655) which consisted of 100 mM Tris-HCl (pH 9.4), 0.15 mg/ml BCIP, 0.3 mg/ml NBT, 5 mM $MgCl_2$ for 30–90 min at 37 °C. After staining, gels were washed with deionized water for 2 h to remove excess substrate and photographed with a digital camera and analyzed using Kodak Digital Science 1D software, version 2.0.2 (Kodak Scientific Imaging Systems).

Binding to Cationized Nylon and Chemiluminescent Detection—The GAG composition of fusion proteins was quantified using a modification of a dot blot assay originally developed for studying molecules with $^{35}SO_4$-labeled GAG chains. DEAE-purified fusion protein samples treated with no enzyme, heparitinase, chondroitinase, or both heparitinase and chondroitinase were blotted individually onto cationic nylon (Zeta-Probe, Bio-Rad) according to the protocol of Rapraeger and Yeaman (23) except that the binding buffer (10 mM Tris-HCl (pH 8.0), 8 M urea, 0.1% Triton X-100, 1 mM Na_2SO_4) contained 1 mM $MgSO_4$ instead of Na_2SO_4. Also, to preserve AP activity, samples were not boiled. After binding, the blot was washed twice in 50 ml of 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM $MgCl_2$, and once in 50 ml of 100 mM diethanolamine (pH 9.8), 1 mM $MgCl_2$. For detection of bound fusion proteins, disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2′-(5′-chloro)tricyclo[3.3.1.1^{3,7}]decan}-4-yl)phenyl (CSPD) AP chemiluminescent substrate solution (Tropix, Foster City, CA) was applied per manufacturer's instructions. The blot was drained, wrapped in plastic, and exposed to HyperFilm ECL (Amersham Pharmacia Biotech). The exposed film was then digitally scanned, and the image was analyzed with NIH Image software. Multiple sample dilutions and exposure times were used to verify that all measurements were taken within the linear range of detection.

Under these blotting conditions, only molecules containing GAG chains should remain associated with the blotting membrane (23). In practice, some AP activity does associate even when fusion proteins that are unglycanated, or have had their GAG chains digested to completion (as confirmed by SDS-PAGE), are examined. Consequently, in determining the reduction in binding of any fusion protein produced by digestion with either heparitinase or chondroitinase, the signal obtained when both enzymes were used together was subtracted from values obtained when a single enzyme or no enzyme was used. After applying this correction, the fraction of any fusion protein that was taken to be glycanated with HS alone was calculated as the fractional decrease in binding produced by heparitinase. The fraction taken to be glycanated with CS alone was calculated as the fractional decrease in binding produced by chondroitinase. The difference, if any, between the sum of these two fractions and unity, was taken to reflect the fraction of molecules glycanated with both HS and CS, *i.e.* hybrid PGs.

Quantification of glycanation by this method depends upon the assumption that the specific activity of AP fusion proteins is not influenced by the presence of nearby GAG chains. This assumption was validated by control experiments in which the specific activities of AP fusion proteins bearing HS (GPC(N1/N2/G)-AP) and CS (GPC(G)-AP) were tested before and after exposure to GAG lyases and were found not to change (data not shown).

Control experiments also supported the assumption that the protein components to which AP is fused also have little if any effect on AP-specific activity. For example, equivalent molar amounts of GPC(N1/N2)-AP and unfused AP (as determined from Coomassie-stained gels) were found to have identical AP activity (data not shown). This observation agrees with experiences of investigators (21, 24, 25) who have fused AP to a variety of proteins.

Selection, Metabolic Labeling, and Analysis of Cell Lines—For transfection and selection of cells stably expressing either AP or GPC(N1/N2)-AP, CHO cells were seeded at a density of 5×10^5 cells in a 100-mm diameter culture dish in complete Ham's F-12 medium. After 1 day, cells were transfected with 5 μg of *APtag-6His*/AP or *APtag-6His*/GPC(N1/N2)-AP with LipofectAMINE in Opti-MEM medium. The medium was changed after 12 h to complete Ham's F-12 medium, and cells were incubated an additional 3 days. Next, the medium was changed to growth medium containing 400 μg/ml G418 and was replaced every 3–4

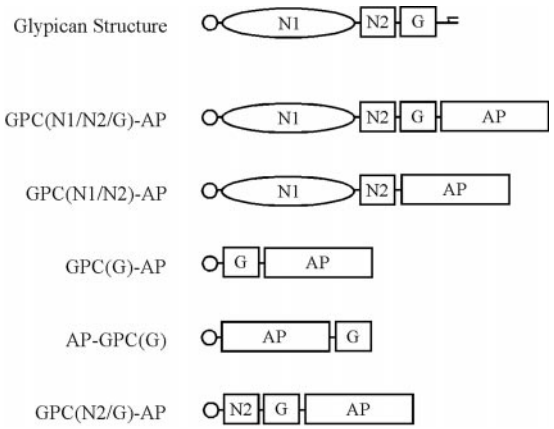

FIG. 1. **Structural representation of glypican-1 and the regions used to produce constructs encoding AP fusion proteins.** All constructs contain the glypican-1 signal peptide (*open circle*) except AP-GPC(G), which contains the native AP signal peptide. On the *far right* of the *top structure* represents the GPI membrane anchor. Tested glypican-1 regions are the N-terminal globular domain (subdivided into regions *N1* and *N2*) and the GAG attachment domain (*G*). All fusion proteins were designed with the endogenous glypican-1 GPI anchorage signal removed to permit efficient secretion into culture medium. All constructs have AP fused to the C terminus except AP-GPC(G).

FIG. 2. **SDS-PAGE of media from transfected COS cells.** Cells were transfected with plasmids encoding GPC(N1/N2/G)-AP (*lane 1*), GPC(N1/N2)-AP (*lane 2*), and GPC(G)-AP (*lane 3*). Fusion proteins (100 fmol per lane) were visualized after electrophoresis by direct AP staining of gels, as described under "Materials and Methods."

days for 2 weeks. Cells were then cloned by limiting dilution in 96-well plates in the continued presence of 200 μg/ml G418. Clones were tested for GPC-AP expression by assaying conditioned media for AP activity. This involved heating 50 μl of conditioned medium to 65 °C for 10 min (to inactivate endogenous phosphatase activity) and mixing with an equal volume of 1 M diethanolamine (pH 9.8), 0.5 mM $MgCl_2$, 0.5 mg/ml bovine serum albumin, 12 mM *p*-nitrophenyl phosphate (Sigma 104–105) for 10 min, and measuring absorbance at 405 nm (26). Colonies with the highest AP expression were split into duplicate 35-mm dishes. At confluence, one of the two duplicate dishes of cells was processed for histochemical detection of AP (24) using BCIP/NBT. Only the clones for which 100% of the cells were expressing AP activity were expanded for further experimentation.

Wild-type CHO cells and cell lines stably expressing either AP (960 fmol/ml/h), a low level of GPC(N1/N2)-AP (584 fmol/ml/h), or a high level of GPC(N1/N2)-AP (1150 fmol/ml/h) were seeded at a density of 6.25×10^4 cells per 35-mm plate and maintained in 2 ml of complete growth medium for 24 h. The concentrations of secreted fusion proteins were calculated from their AP activity, as described (21). For metabolic labeling, cells were first washed twice with 5 ml of phosphate-buffered saline and then incubated in 1 ml of sulfate-deficient medium containing 1 mCi of $Na_2{}^{35}SO_4$ (ICN, Costa Mesa, CA) which had been equilibrated in a 37 °C, 5% CO_2 environment. The sulfate-deficient medium consisted of Ham's F-12 medium with chloride salts replacing all sulfate salts (27) supplemented with 7.5% (v/v) fetal bovine serum that had been exhaustively dialyzed against phosphate-buffered saline. After 1 h of incubation, media were harvested, and the secreted PGs were subsequently purified via DEAE-Sepharose chromatography (20). Eluates were subjected to liquid scintillation counting to determine total $^{35}SO_4$ incorporation into GAGs, which did not differ significantly among the cell lines.

GAG chains were released from purified PGs via alkaline-borohydride treatment followed by ethanol precipitation (28). Compositional analysis of free GAG chains was performed using the original dot blot procedure of Rapraeger and Yeaman (23), either with or without prior digestion with 1 unit/ml chondroitinase ABC (Seikagaku) at 37 °C for 1 or 2 h (values at both time points were not significantly different, so both types of data were pooled). The amount of radioactivity bound after chondroitinase treatment as a percentage of the undigested sample was taken to represent the percentage of HS; this approach was justified by SDS-PAGE analyses (29) indicating a lack of detectable [^{35}S]sulfate labeling of any macromolecules other than CS and HS. Parallel experiments that examined the GAG composition of cell-associated, rather than secreted, PGs revealed that the former consisted almost entirely of HSPGs. Not surprisingly, the relative content of cell-associated HS did not change in any of the experimental conditions (data not shown).

RESULTS

Glypican core proteins share a common structure in which a highly conserved cysteine-rich N-terminal domain is followed by a smaller region with multiple Ser-Gly dipeptide sequences and finally by a short, poorly conserved C-terminal region that provides signals for GPI anchorage. In glypican-1, the second of these regions contains three consecutive Ser-Gly dipeptides, and mutagenesis experiments have shown that GAGs are attached to glypican-1 only at these serine residues (30). The N-terminal domain has no known function but is the site of substantial intrachain disulfide bonding, consistent with folding into a compact globular structure (20, 31). We may thus divide mature rat glypican-1 into a globular domain, a GAG attachment domain, and a GPI anchor (Fig. 1). In this study we have further subdivided the N-terminal globular domain into an initial portion that contains all 14 cysteine residues (region N1, 392 amino acids following removal of the signal peptide), followed by a cysteine-free segment (region N2, 58 amino acids) that separates it from the GAG attachment domain (region G, 56 amino acids, of which Ser-Gly dipeptides represent the 13th through the 18th). The boundary between region N2 and region G (after amino acid Gly^{473}) occurs just prior to a cluster of acidic amino acids that is present in all glypicans.

To investigate how sequences in PG core proteins influence the types of GAG chains they bear, one must be able to quantify the numbers and types of GAG chains that are added to experimental core protein molecules. Typically, GAG biosynthesis is measured through the use of metabolic labels, such as [^{3}H]glucosamine and [^{35}S]sulfate. We chose instead to place a tag on the core protein and use a combination of chromatographic behavior and enzyme susceptibility to distinguish PG species that are glycanated in different ways.

Briefly, various regions of glypican-1 were genetically fused at their C terminus to a secreted form of AP. In one case AP was used as an N-terminal tag rather than a C-terminal tag. In all cases the glypican-1 GAG attachment domain was truncated to remove its GPI anchorage signal, so that all constructs encoded secreted rather than membrane-anchored proteins. DNA fragments encoding fusion proteins, as illustrated in Fig. 1, were

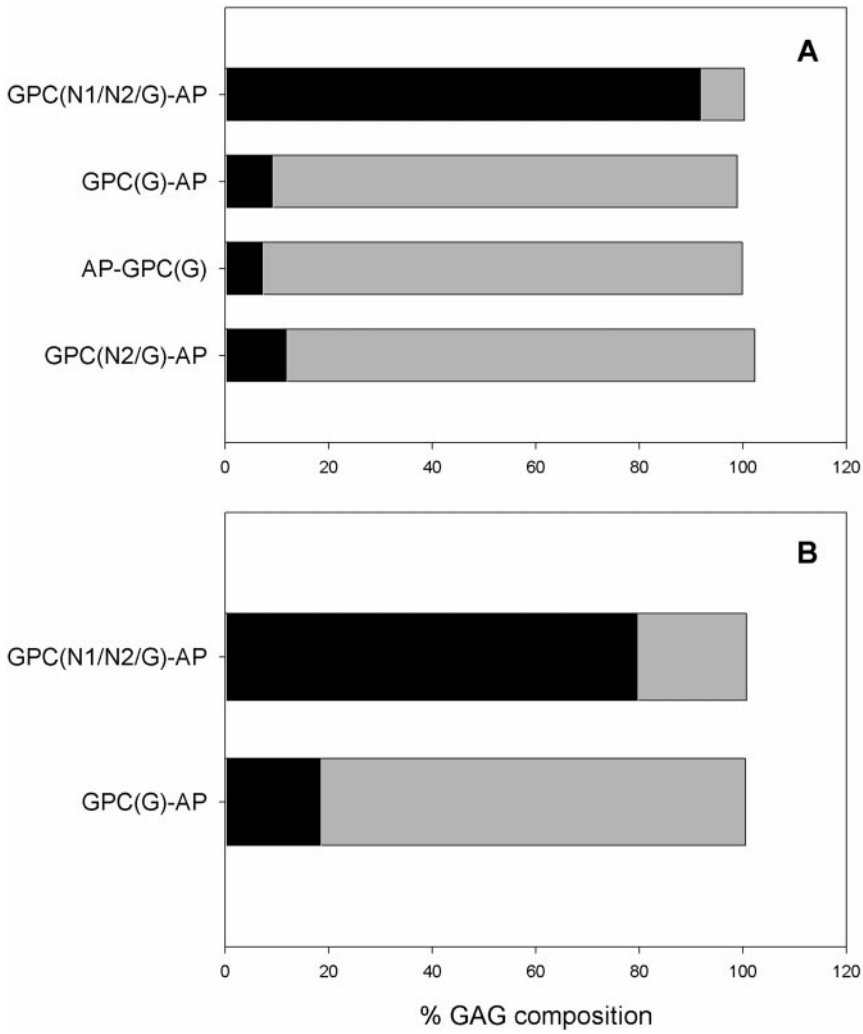

FIG. 3. **GAG composition of fusion proteins containing various regions of glypican-1.** COS (*A*) and CHO (*B*) cells were transfected with plasmids encoding the indicated fusion proteins. Glycanated fusion proteins were purified from conditioned media by DEAE chromatography and analyzed for GAG composition (see "Materials and Methods"). *Solid bars* correspond to the percentage of DEAE-eluted fusion proteins that bore HS, and *gray bars* correspond to the percent that bore CS. Measurements of HSPGs and CSPGs were made independently.

cloned into the mammalian expression vector pcDNA3 and subsequently transfected into COS or CHO cells. After transfection, culture medium was harvested after 3–5 days and analyzed directly by SDS-PAGE. By omitting the boiling of samples prior to electrophoresis, it was possible to retain sufficient AP activity so that the migration of fusion proteins could be visualized by soaking gels in a colorimetric AP substrate solution.

As shown in Fig. 2, a fusion protein containing both the glypican-1 globular domain and the GAG attachment domain (GPC(N1/N2/G)-AP) and one containing just the GAG attachment domain (GPC(G)-AP) both migrated as broad smears on SDS-PAGE, consistent with the presence of GAG chains. In contrast, GPC(N1/N2)-AP, which possesses only the globular domain, migrated as a sharp band, consistent with a non-glycosylated or a poorly glycosylated protein. These results demonstrate the utility of AP tagging for rapid assessment of GAG addition to fusion proteins and support the conclusions of Mertens *et al.* (30) that only the Ser-Gly dipeptides in region G of glypican-1 support GAG biosynthesis.

Requirement of the Globular Domain for Preferential HS Assembly—To determine whether the smears in Fig. 2 reflected the presence of HS or CS chains, PGs were purified by anion exchange chromatography from the conditioned media of transfected cells and were subjected to enzymatic digestion with either heparitinase (heparinase III) or chondroitinase ABC. Products were analyzed either by SDS-PAGE as in Fig. 2 or, for more quantitative measurements, by dot blotting onto cationized nylon, followed by determination of bound AP activity (see "Materials and Methods").

As shown in Fig. 3*A*, such analysis revealed that GPC(N1/N2/G)-AP, when expressed in COS cells, was made almost exclusively as HSPGs (92%) with the balance consisting of CSPGs. Native glypican-1 that has been isolated from a variety of sources (14–16) has been reported to bear solely HS. It is not clear whether the small proportion of CSPGs observed in the present study also occurs with native glypican-1 (but has gone unnoticed) or whether it is a consequence of the experimental methods (*e.g.* overexpression, removal of the GPI anchor, addition of AP). Nonetheless, it appears that the GAG preference of the GPC(N1/N2/G)-AP fusion protein is reasonably close to that of native glypican-1.

In contrast to GPC(N1/N2/G)-AP, a fusion protein with the globular domain deleted, GPC(G)-AP, was produced in COS cells at similar levels but consisted almost entirely of CSPGs (90%) with the balance consisting of HSPGs (Fig. 3*A*). The loss of preferential HS assembly caused by globular domain deletion was not cell type-specific, as similar results were generated with the same fusion proteins expressed in CHO cells (Fig. 3*B*). Thus, whereas the glypican-1 GAG attachment domain contains sequences necessary for the attachment of GAGs, the globular domain contains information that dramatically alters GAG preference from CS to HS.

The dependence of HS glycanation on the presence of the globular domain could reflect specific sequence information in that domain, or it could simply reflect a requirement that the

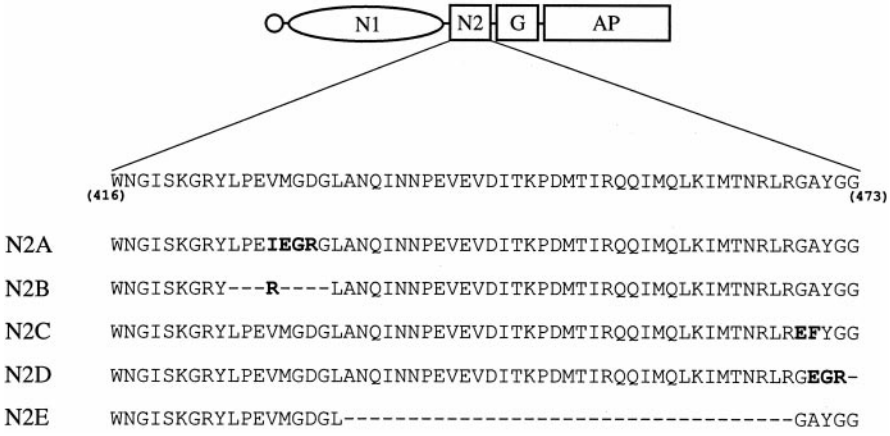

FIG. 4. **Region N2 mutants.** Mutations were introduced into region N2 of the plasmid encoding GPC(N1/N2/G)-AP to produce five mutants, *N2A* to **N2E**. Amino acid sequence in region N2 is shown in *single-letter code*, beginning with ^{416}W and ending with ^{473}G. *Letters* in *bold* denote substitutions, and *dashes* represent deletions.

GAG attachment site not be located close to the N terminus of the protein. To distinguish between these possibilities, we generated a fusion protein in which AP was fused to the N terminus, rather than the C terminus, of region G. AP is a globular protein with a mass of 67 kDa; thus it is not unlike the glypican-1 globular domain itself, which has a mass of 58 kDa. When expressed in COS cells, this construct, AP-GPC(G), had a GAG composition almost identical to that of GPC(G)-AP (Fig. 3*A*). Thus, substitution of the glypican-1 globular domain with a heterologous globular domain of similar mass failed to restore preferential HS assembly.

We next attempted to determine whether the ability of the glypican-1 globular domain to promote HS assembly could be localized to sequences most closely adjacent to the GAG attachment site, by testing GPC(N2/G)-AP, which includes the last 58 amino acids of the globular domain. As shown in Fig. 3*A*, only 12% of that fusion protein was made as HSPGs, with the remainder consisting of CSPGs, a composition not significantly different from that of GPC(G)-AP. Thus, region N2 is not sufficient to promote HS assembly.

To test whether region N2 is necessary for promotion of HS assembly, mutations were introduced into the region in the context of the full-length fusion protein construct GPC(N1/N2/G)-AP (Fig. 4). Mutations were located either toward the N-terminal end (N2A and N2B) or near the C-terminal end (N2C and N2D) of region N2. Of the glycanated fusion proteins produced by these constructs, 74 and 78% of N2A and N2B, respectively, were made as HSPGs (Fig. 5). In the case of the more C-terminal mutations, 45 and 47% of N2C and N2D, respectively, were expressed as HSPGs. In all 4 cases, the remaining glycanated fusion proteins were made as CSPGs. Together, these results indicate that structures in both region N1 and N2 are required for ensuring preferential HS assembly on glypican-1.

Finally, we sought to determine, with mutant N2E (Fig. 4), whether region N1, independent of region N2, could promote HS assembly. This mutant has 60% (35 out of 58 amino acids) of region N2 deleted (Fig. 4). However, no AP activity was detected in media from transfected cells, suggesting that this fusion protein was either not expressed or not properly folded.

The Glypican-1 GAG Attachment Domain Does Not Support HS/CS Hybrid Formation—Since each of the three serine residues in the glypican-1 GAG attachment domain can prime a GAG chain (30), in principle both types of chains, HS and CS, could be assembled at the same attachment domain, producing HS/CS hybrid PGs. In all of the experiments described above (Fig. 3 and 5), we found that formation of hybrid PGs (calculated as the proportion of core proteins that remained able to bind cationized nylon after treatment with either heparitinase or chondroitinase but not after treatment with both enzymes)

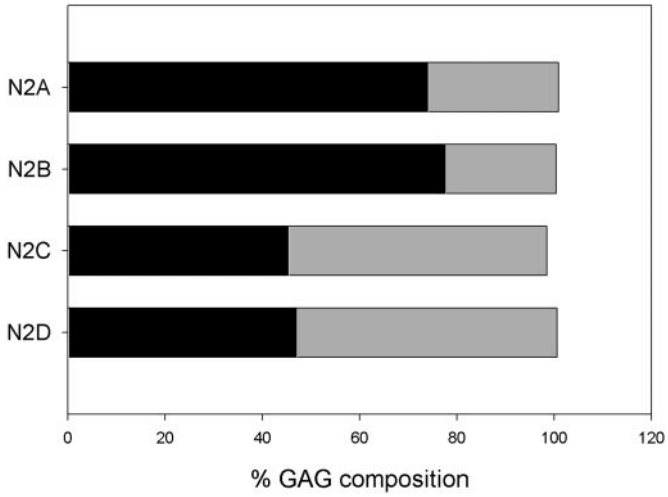

FIG. 5. **GAG composition of region N2 mutants.** Cells were transfected with plasmids encoding GPC(N1/N2/G)-AP containing various mutations in region N2. Four mutants, N2A, N2B, N2C, and N2D, were expressed and subsequently purified from the conditioned media and analyzed for GAG composition (see "Materials and Methods"). Cells transfected with N2E did not express detectable AP activity either in the medium or within cells. *Solid bars* correspond to the percentage of glycanated fusion proteins that bore HS, and *gray bars* correspond to the percent that bore CS. Measurements of HSPGs and CSPGs were made independently.

was below the limits of detection. These results suggest that the three GAG attachment sites in glypican-1 are functionally linked, so that either HS or CS may be primed but not both.

The Globular Domain Enhances HS Synthesis on Heterologous GAG Attachment Domains—The N-terminal globular domain of glypican-1 is similar in sequence among all known glypicans but is unrelated to any other known protein domain. This raises the question of whether its ability to enhance HS assembly is limited to GAG attachment domains derived from glypicans. This question was addressed by substituting a GAG attachment site from betaglycan. It has previously been shown that this site can support the attachment of a single chain of either HS or CS (11). AP was fused to this domain at its C terminus, either with or without the glypican-1 globular domain at its N terminus, thus producing GPC(N1/N2)-B-AP and B-AP, respectively (Fig. 6*A*). When these constructs were transfected into COS cells, the results showed that the globular domain increased HS assembly on the betaglycan GAG attachment domain, from 8.1 to 55% (Fig. 6*B*). Thus, the glypican-1 globular domain can influence a heterologous GAG attachment domain, even one that possesses only a single GAG attachment site.

As a more stringent test of the activity of the glypican-1

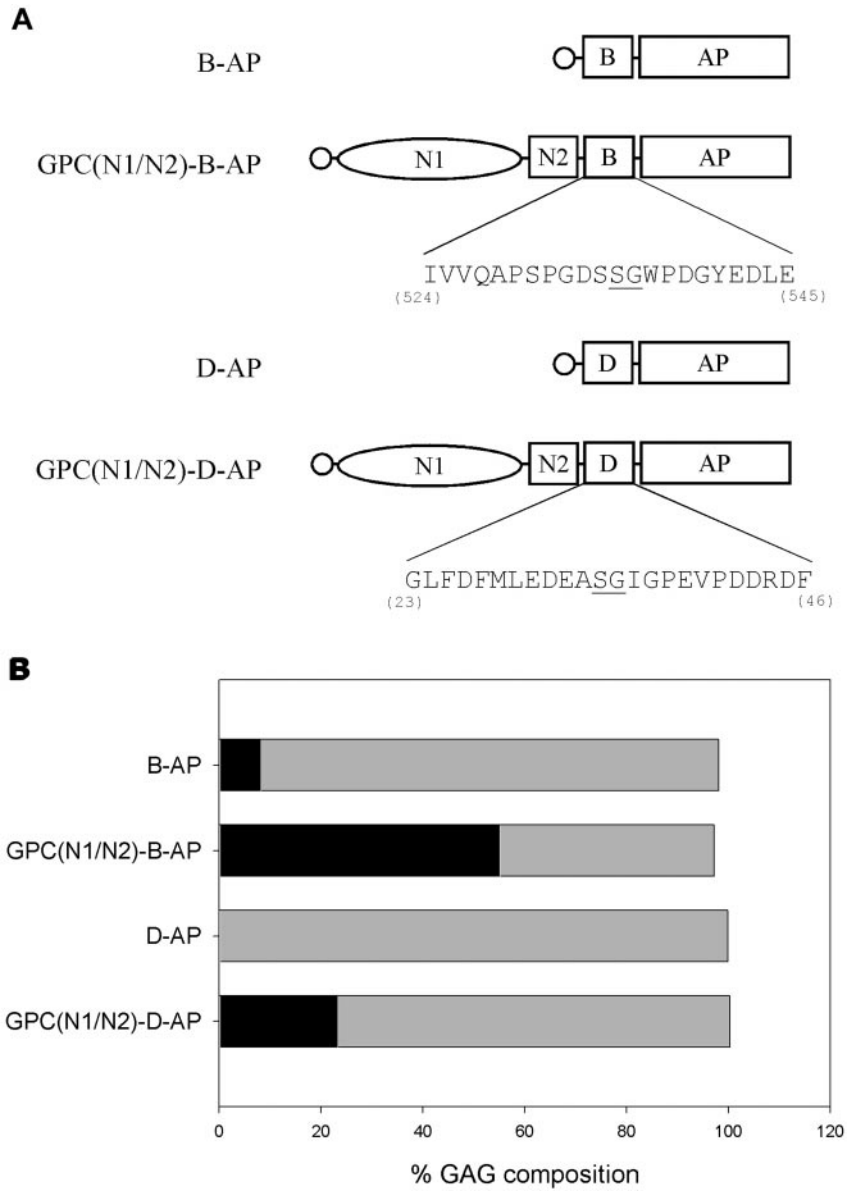

FIG. 6. **Effect of the glypican-1 globular domain on heterologous GAG attachment domains.** *A,* plasmids encoding AP fusion proteins were made using either the first GAG attachment domain from betaglycan (*B*) and the GAG attachment domain from decorin (*D*). The *numbers* in *parentheses* indicate the positions of the amino acids relative to the start methionine of the native PGs. Both types of constructs were made either without (*B-AP* and *D-AP*) or with (*GPC(N1/N2)-B-AP* and *GPC(N1/N2)-D-AP*) regions N1 and N2 of glypican-1 fused to the N terminus. All constructs contained the signal peptide of glypican-1 (*open circle*). Serine-glycine dipeptides in the GAG attachment regions are *underlined*. *B,* cells were transfected with the plasmids depicted in *A*. Glycanated fusion proteins were purified from the conditioned media and analyzed for GAG composition (see "Materials and Methods"). *Solid bars* correspond to the percentage of glycanated fusion proteins that bore HS, and *gray bars* correspond to the percent that bore CS. Measurements of HSPGs and CSPGs were made independently.

globular domain, we carried out an analogous set of experiments in which the GAG attachment domain was derived from decorin (Fig. 6*A*). Decorin carries a single CS chain at this site, and has never been observed to carry HS (32–34). Indeed, in our hands the decorin GAG attachment domain, when expressed as an AP fusion protein, carried only CS (Fig. 6*B*). In contrast, when the glypican-1 globular domain was fused to the N terminus of the same construct, the GAG attachment domain supported 23% HS. Thus, the glypican-1 globular domain can induce HS assembly even in an environment that is otherwise nonpermissive for the synthesis of HS.

The Glypican-1 Globular Domain Decreases Overall Glycanation—As mentioned earlier, the core proteins of many naturally occurring PGs may be produced by cells in an unglycanated form. Studies by Sugahara and colleagues (10) suggest that glycanation is blocked at the level of the α-GlcNAc- and β-GalNAc-T1 enzymes, the same enzymes that determine whether a chain becomes HS or CS. Consequently, enzymatic action or inaction at this step may be viewed as producing any of three mutually exclusive outcomes as follows: glycanation with HS, CS, or no GAG.

To gain greater insight into the mechanism by which the glypican-1 globular domain influences glycanation, it would therefore be advantageous to quantify not only the relative proportions of HS and CS glycanation of fusion proteins but also to track the molecules that are unglycanated. Fortunately, the AP fusion protein method is particularly suited to this task. The unglycanated fraction can be determined from the amount of AP enzymatic activity that fails to bind DEAE-Sepharose even when not digested with any GAG lyase. Results can be (and were) confirmed by SDS-PAGE, in which unglycanated material noticeably migrates as a sharp band.

When such measurements were made for the fusion proteins described above, it was first noted that different GAG attachment domains supported very different levels of glycanation, despite similar levels of protein expression. For example, fusion proteins containing glypican-1, betaglycan, and decorin GAG attachment domains (without any globular domain) exhibited 90, 12, and 0.55% glycanation, respectively, in COS cells (Fig. 7). To our knowledge this is the first demonstration that the GAG attachment domains of native PGs can vary so dramatically in their ability to support GAG biosynthesis in a single cell type.

We also observed that inclusion of the glypican-1 globular domain in fusion proteins caused, in every case, a decrease in the fraction of material that was glycanated (Fig. 7). For ex-

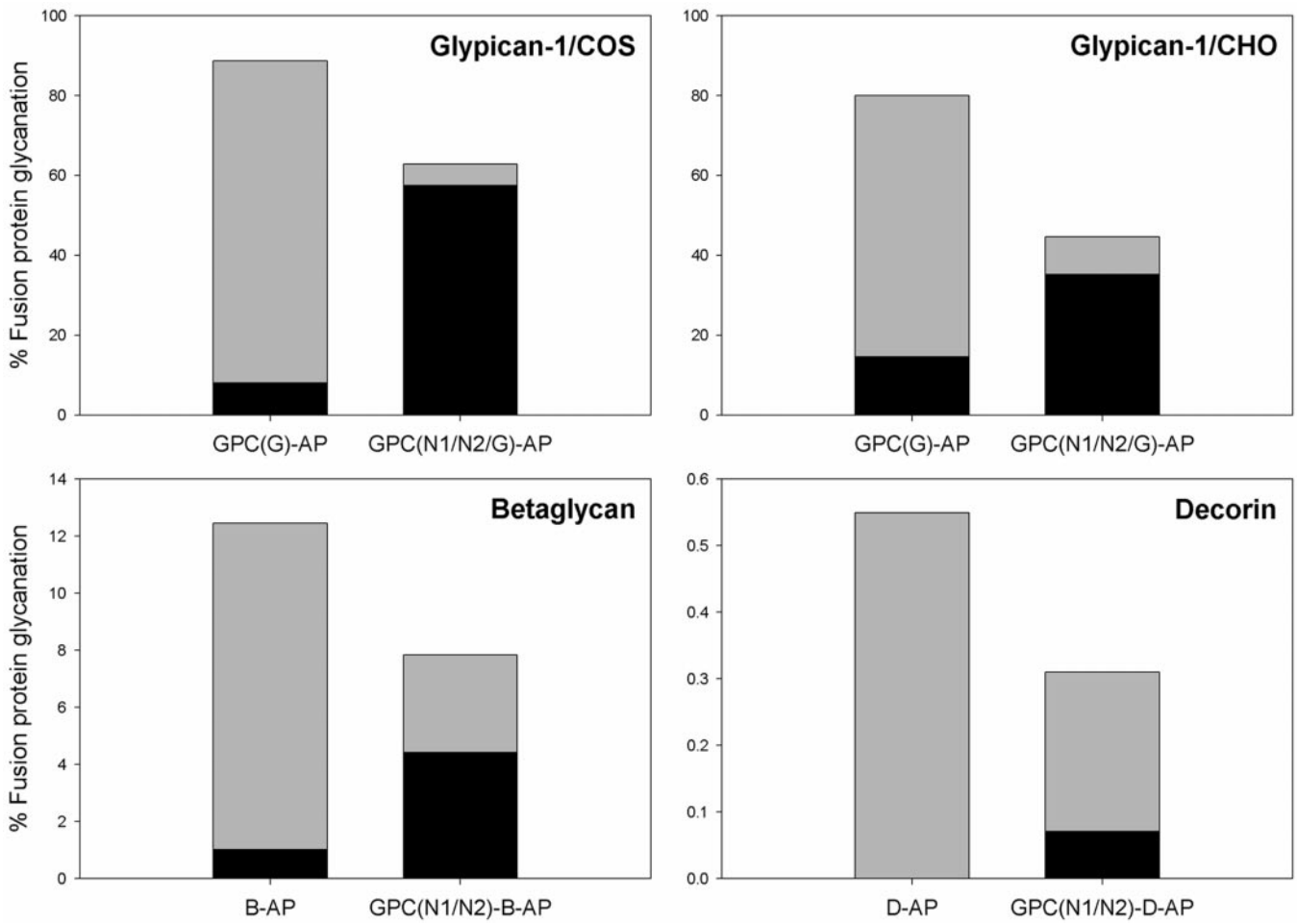

FIG. 7. **Effect of the glypican-1 globular domain on the efficiency of glycanation of fusion proteins.** Measurements of retention of fusion proteins by DEAE-Sepharose, which reveals the fraction of molecules that is glycanated, are plotted together with measurements of the proportions of molecules bearing HS (*solid bars*) *versus* CS (*gray bars*). *Upper left, GPC(G)-AP* and *GPC(N1/N2/G)-AP* expressed in COS cells. *Upper right, GPC(G)-AP* and *GPC(N1/N2/G)-AP* expressed in CHO cells. *Lower left, B-AP* and *GPC(N1/N2)-B-AP* expressed in COS cells, and *lower right, D-AP* and *GPC(N1/N2)-D-AP* expressed in COS cells.

ample, in COS cells, a fusion protein containing the glypican-1 GAG attachment domain but lacking the glypican-1 globular domain was 90% glycanated, whereas one containing both the globular and GAG attachment domains was only 63% glycanated. In CHO cells the values for the same molecules were 80 and 44%, respectively. Similar reductions were observed for fusion proteins bearing betaglycan- and decorin-derived GAG attachment domains (Fig. 7).

From these results it may be concluded that the glypican-1 globular domain acts both to decrease the probability that a core protein will become glycanated and yet increase the probability that glycanated molecules will possess HS chains, rather than CS chains. The mechanistic implications of these observations will be discussed below.

Effects of the Glypican-1 Globular Domain on Total Cellular HS Assembly—The results presented above demonstrate that the glypican-1 globular domain can act in *cis, i.e.* when directly attached to a GAG attachment domain, to promote HS assembly. We also considered that possibility that the globular domain might act in *trans*, *i.e.* on PG molecules to which it is not attached. Initial evidence that this might be the case was obtained by cotransfecting COS cells with vectors encoding GPC(G)-AP along with either GPC(N1/N2)-AP or AP alone (Fig. 8). The aim of the experiment was to test whether the pattern of glycanation of GPC(G)-AP could be altered simply by expressing the other constructs together with it. Since GPC(N1/N2)-AP and AP cannot be glycanated, the fact that they possess AP activity does not interfere with the determination of relative proportion of HS and CS on GPC(G)-AP.

As shown in Fig. 8, when GPC(G)-AP was coexpressed with AP, 12% of it was made as HSPGs with the remainder CSPGs; these values are similar to those obtained when GPC(G)-AP was expressed alone (Fig. 3*A*). In contrast, when GPC(G)-AP was coexpressed with GPC(N1/N2)-AP, 36% of it was produced as HSPGs, with the balance being CSPGs (Fig. 8). Thus, the globular domain acted to enhance HS assembly on molecules where it was not attached.

If the mechanism of action of the glypican-1 globular domain acting in *trans* is the same as that used when acting in *cis*, one would predict that its influence should not be limited to fusion proteins with glypican-derived GAG attachment domains but also extend to other GAG attachment domains, including those found on the endogenous PGs of the cell. We thus generated CHO cell lines stably expressing GPC(N1/N2)-AP (two clones with expression levels of 580 fmol/ml/h (clone 1) and 1150 fmol/ml/h (clone 2)) as well as a control clone expressing a comparable amount of AP alone (960 fmol/ml/h), and we compared the GAG compositions of endogenous PGs among these as well as wild-type CHO cells. Briefly, cells were metabolically labeled for 1 h with $^{35}SO_4$, and GAGs isolated from secreted PGs were analyzed according to the original procedure of Rapraeger and Yeaman (23). The results, shown in Fig. 9, indicate that the proportion of HS synthesis by CHO cells was significantly increased in response to the expression of GPC(N1/

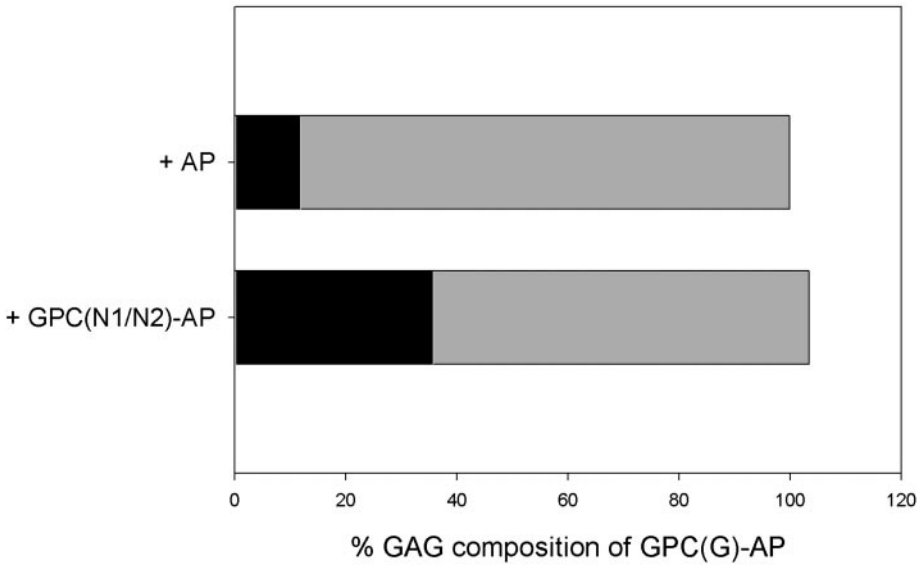

FIG. 8. **Effect of coexpression of the glypican-1 globular domain with GPC(G)-AP.** Cells were transfected with a plasmid encoding GPC(G)-AP together with a 10-fold excess of either *AP* (*upper bar*) or *GPC(N1/N2)-AP* (*lower bar*). The glycanated material (*i.e.* only that encoded by GPC(G)-AP) was purified from the conditioned media and analyzed for GAG composition (see "Materials and Methods"). *Solid bars* correspond to the percentage of glycanated fusion proteins that bore HS, and *gray bars* correspond to the percent that bore CS. Measurements of HSPGs and CSPGs were made independently.

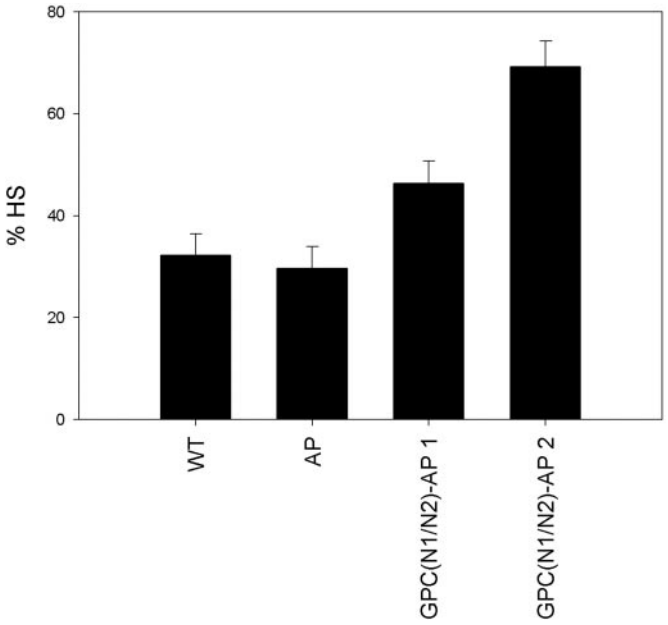

FIG. 9. **Effect of the glypican-1 globular domain on the GAG composition of endogenous cellular PGs.** CHO cells that were either untransfected (wild-type, *WT*) or were transfected with and stably expressed AP at 960 fmol/ml/h, GPC(N1/N2)-AP at 584 fmol/ml/h (clone 1), or GPC(N1/N2)-AP at 1150 fmol/ml/h (clone 2) were metabolically labeled with $Na_2{}^{35}SO_4$ for 1 h at 37 °C. The conditioned media were processed, and the GAG composition was determined as described under "Materials and Methods." The data are averages (±S.D.) from three separate repetitions of the metabolic labeling. Clones 1 and 2 both showed a significant increase in the proportion of GAG that was HS ($p < 0.05$), whereas the clone expressing AP was not significantly different from the untransfected control.

N2)-AP but was not affected by expression of AP alone. Furthermore, the greatest increase in HS synthesis occurred in the clone with the highest level of GPC(N1/N2)-AP expression. Thus, expression of the glypican-1 globular domain in a cell line can have profound and global effects on cellular GAG synthesis.

DISCUSSION

The results presented above imply that the globular domain of glypican-1 plays an essential role in directing the preferential synthesis of HS, rather than CS, on the glypican-1 core protein. Deletion of this domain, deletion of its N-terminal cysteine-rich region (region N1), or alteration of amino acid sequences in its C-terminal cysteine-free segment (region N2), all caused a decrease in the proportion of glycanated glypican-1 fusion proteins that carried HS. Although these results were obtained using secreted AP fusion proteins, rather than native PGs, it is noteworthy that a fusion protein containing nearly the entire glypican-1 core protein sequence (*i.e.* GPC(N1/N2/G)-AP) was glycanated much like native glypican-1, *i.e.* almost exclusively with HS. Moreover, the high proportion of CS observed on the AP fusion protein containing just the glypican-1 GAG attachment domain (GPC(G)-AP) was similar to what Zhang *et al.* (12) observed for protein A fusions of approximately the same domain of human glypican-1. Thus, there is little reason to believe that the use of AP to tag PG core proteins adversely affects the way such proteins are handled by the GAG biosynthetic machinery. Furthermore, although the various fusion protein constructs used in this study were expressed at a variety of different levels (mostly within a factor of 1.5 of each other), relative expression levels did not correlate in any systematic fashion with whether a fusion protein bore predominantly HS or CS (Table I). Thus, the differences in glycanation among fusion proteins in the present study are unlikely to be an indirect consequence of variable saturation of rate-limiting step(s) in the GAG biosynthetic machinery. Rather, they appear to reflect direct effects of the glypican-1 globular domain on the biosynthetic process.

The Glypican-1 Globular Domain Is an Independent and General Enhancer of HS Glycanation—The results of fusing the glypican-1 globular domain to the GAG attachment domains of the heterologous PG core proteins betaglycan and decorin show that the globular domain can promote the attachment of HS chains, even to a site (derived from decorin) that otherwise carries undetectable levels of HS (both in fusion protein form and in the native PG). In general, the glypican-1 globular domain has a more profound effect on GAG choice than any other established domain in a PG core protein (outside of the amino acid residues immediately surrounding sites of GAG attachment). For example, in perlecan, a "SEA" domain 20 amino acids C-terminal to a GAG attachment site has been shown to elevate modestly the proportion of HS, from 61–62 to 73–81% (13).

Although the glypican-1 globular domain strongly promoted the attachment of HS to various GAG attachment domains, the proportion of HS observed in such fusion proteins varied from construct to construct. The highest proportion of HS glycanation (92%) was observed on the GAG attachment domain from

TABLE I
Measured parameters for all fusion proteins

Data are summarized for all of the AP fusion proteins described in Figs. 3–7. "Total protein" refers to the level of expression of each fusion protein at the time of harvest of conditioned medium and was calculated from the AP activity (see "Materials and Methods"). The proportions of molecules bearing HS, CS, and no GAG were determined as described in the text. Although differences in the expression level of different constructs are noticeable, these did not correlate significantly with degree or type of glycanation. ND, not determined.

Fusion protein	Total protein	PG	HSPG	CSPG
	pmol/ml	*% total protein*	*% PG*	
GPC (N1/N2)-AP	61	ND	ND	ND
GPC (N1/N2/G)-AP	31	63	92	8.5
GPC (G)-AP	44	90	9.1	90
AP-GPC (G)	46	87	7.2	93
GPC (N2/G)-AP	33	83	12	90
GPC (N1/N2/G)-AP, CHO	17	44	80	21
GPC (G)-AP, CHO	23	80	18	82
N2A	7.5	59	74	27
N2B	7.4	51	78	23
N2C	20	68	45	53
N2D	17	82	47	54
B-AP	34	12	8.1	90
GPC (N1/N2)-B-AP	57	8.1	55	42
D-AP	40	0.55	<0.01	>99
GPC (N1/N2)-D-AP	39	0.31	23	77

glypican-1 (which predominantly bears HS *in vivo*); an intermediate level (55% HS) was seen on a betaglycan GAG attachment domain (which bears either HS or CS *in vivo*), and the lowest level (23% HS) was observed on the GAG attachment domain of decorin (which bears only CS *in vivo*). Thus, the influences of sequences local to the GAG attachment site of PGs are still functional even in the presence of the glypican-1 globular domain, suggesting that local and distant sequence elements in a PG core protein act through independent, potentially additive mechanisms.

Mechanism of Action of the Glypican-1 Globular Domain—In principle, there are numerous ways in which the glypican-1 globular domain might influence the glycanation of PG core proteins. However, its ability to affect core proteins to which it is not physically attached (*i.e.* in *trans* (Figs. 8 and 9)) would seem to rule out models based simply on steric hindrance of the access of core proteins to glycosyltransferases. Instead, it would appear that the globular domain must interact with some component of the biosynthetic pathway, *e.g.* a glycosyltransferase, a cofactor required for the activity or localization of such an enzyme, or some limiting component required for trafficking PG core proteins within the Golgi apparatus. Although the data do not indicate what molecule the glypican-1 globular domain interacts, it is possible to narrow the choices if we assume that a common mechanism underlies both the ability of the globular domain to increase the proportion of PGs bearing HS (Figs. 3–6) and the consistent ability of the globular domain to decrease the proportion of molecules bearing GAG at all (Fig. 7).

For example, in transfection experiments involving COS cells (Fig. 7), construct GPC(G)-AP yielded a ratio of HS:CS:no GAG of ~10:80:10. When the glypican-1 globular domain was present (*i.e.* GPC(N1/N2/G)-AP), the ratio was 58:5:37. It should be evident that this increase in HS cannot be accounted for by the recruitment of molecules solely from the unglycanated pool (as the latter is not sufficiently large). Rather, molecules that otherwise would have become CSPGs are bearing HS instead. In principle, this could be due either to an increase in the activity of a biosynthetic pathway leading to HS or a decrease in the activity of a pathway leading to CS. Yet, when we take into account the decrease in overall glycanation that is also brought about by the globular domain, it is clear that an increase in HS synthesis could only explain the data if it were accompanied by a parallel, even greater decrease in CS synthesis. In contrast, inhibition of CS synthesis alone would bring about both an increase in the proportion of molecules bearing HS and a decrease in the proportion of molecules bearing GAG at all, as long as those molecules that are diverted away from CS synthesis are not all glycanated with HS.

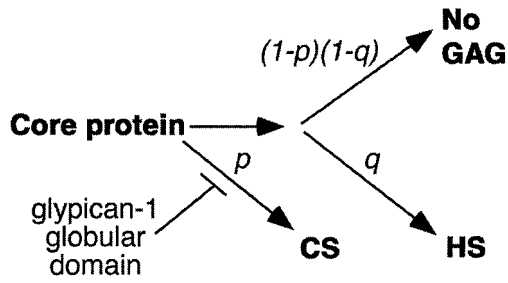

FIG. 10. **Proposed action of the glypican-1 globular domain.** A two-step biosynthetic model is proposed in which the decisions to attach CS or HS to a core protein are made sequentially. Based on the model, the ratios of HS:CS:no GAG for a core protein will be $q(1-p):p:(1-q)(1-p)$, where p and q are the probabilities of initiating CS and HS synthesis at the indicated steps. From observed ratios (*e.g.* Fig. 7), it is possible to solve for values of p and q. Such an exercise reveals that, for experiments comparing GPC(G)-AP with GPC(N1/N2/G)-AP, the effect of the glypican-1 globular domain is to decrease the calculated value of p from ~0.8 to 0.05 (COS cells) or 0.65–0.09 (CHO cells). In contrast, q is little affected (changing from ~0.49–0.61 in COS cells, or 0.43–0.39 in CHO cells). If the model is altered so that the decision to make either HS or no GAG precedes the decision to make CS, it is impossible to fit the data without requiring large changes in both p and q (calculations not shown). Although this model fits the data in Fig. 7 for constructs containing the glypican-1 GAG attachment domain, in the case of the betaglycan and decorin GAG attachment domains, the very large fraction of molecules that are not glycanated requires that additional modifications to the model be made (*e.g.* having some fraction of molecules bypass the GAG attachment pathway altogether).

Fig. 10 develops this idea in a quantitative manner, showing that for both COS and CHO cells, the data from comparing GPC(G)-AP with GPC(N1/N2/G)-AP can be reasonably fit by a biosynthetic model involving two steps. In the first step, some core protein molecules become committed to bear CS, whereas in the second step the remaining molecules are selected either to bear HS or stay unglycanated. The effect of the glypican-1 globular domain is to inhibit the first step, lowering the likelihood of commitment to CS from 65–80 to ~5–9% (see legend to Fig. 10). The numerical agreement between the model and the data suggests that the glypican-1 globular domain may indeed act primarily to inhibit CS synthesis, doing so at a step that precedes commitment to bear, or not bear, HS.

It is tempting to propose that the glypican-1 globular domain

acts at the level of β-GalNAc-TI, the enzyme that actually commits a nascent GAG chain to become CS (5). However, it is also possible that the globular domain acts at the level of protein trafficking, to influence the delivery of molecules to whatever Golgi compartment contains β-GalNAc-TI. Ultimately, answers to these questions may come from identification of molecules with which the glypican-1 globular domain physically interacts (of which none are currently known). To this end, it would be helpful to pursue more systematic mutagenesis to map more precisely the parts of the glypican-1 globular domain that are involved in influencing glycanation. Unfortunately, several attempts to alter sequence in, or truncate, the N1 region of the glypican-1 globular domain have yielded proteins that fail to express, presumably due to misfolding.[3] Undoubtedly, elucidation of the tertiary structure of the highly disulfide-linked globular domain should improve the chances for success of such experiments. In the meantime, it may also be possible to obtain some answers by comparing the globular domains of different glypican-1 family members. For example, when expressed in COS cells, glypican-5 possessed both HS and CS (35), raising the possibility that the glypican-5 globular domain may not be as efficient at influencing the biosynthetic machinery as the glypican-1 domain.

In Vivo Roles of Glypican Globular Domains—It is curious that most of the local sequence elements within a GAG attachment domain that have been shown to enhance glycanation with HS, clusters of acidic amino acids, repetitive serine-glycine dipeptides (11, 12), are particularly highly represented in the GAG attachment domain of glypican-1, yet in the absence of the globular domain it is substituted overwhelmingly with CS. Indeed, it is clear from the data of Zhang *et al.* (12) that although similar local sequence elements are present in the GAG attachment sites of most HSPGs (and lacking in most CSPGs), the isolated attachment sites of HSPGs prime HS to a degree that is both incomplete and highly variable (7–64%, depending on the particular GAG attachment site). The simple interpretation of such data is that the biosynthetic machinery that initiates an HS chain is highly sensitive to subtle sequence differences, such as exact numbers and positions of acidic residues, interference from other amino acids, etc.

In light of the present data, however, we suggest an alternative explanation that most or all of the native HSPG GAG attachment sites studied by Zhang and co-workers (11, 12) are highly and perhaps equally suitable for initiation of HS, but they may vary in their susceptibility to the initiation of CS chains (due to the presence of sequence elements yet undiscovered). This interpretation is particularly reasonable if the model in Fig. 10 is correct, since that model predicts that, even when susceptibility to initiation of HS is maximized (*i.e.* $q = 1$ in the model), differences in susceptibility to CS initiation will dramatically alter the proportion of molecules that ultimately bear HS. Interestingly, it may be possible to support or refute this interpretation by examining the degree of overall glycanation supported by the various GAG attachment sites studies by Zhang and co-workers (11, 12). The prediction made by the model in Fig. 10 is that those molecules that bear the least HS will exhibit the highest level of glycanation.

Whatever the explanation for the inefficient synthesis of HS on the GAG attachment domains of many HSPGs, in the case of glypican-1 efficient glycanation with HS clearly relies heavily on the presence of the globular domain. How the same effect is achieved for other families of HSPGs remains, in most cases, to be worked out. Interestingly, the ability of an exogenous glypican-1 globular domain to influence glycanation in *trans* (Figs. 8 and 9) naturally leads to the speculation that endogenous glypicans might normally exert the same effect on the glycanation of endogenous PGs. Whether that actually happens *in vivo* or whether the *trans* effect is something that occurs only when glypicans are expressed at levels outside the physiological range remains to be evaluated by appropriate means. Ultimately, analysis of patterns of glycanation in the tissues of animals mutant for glypican core protein genes (*e.g.* see Refs. 36 and 37) might shed some light on this issue.

Acknowledgments—We thank Jeff Esko for helpful advice and for the gift of reagents for metabolic labeling of cells. We also thank Paul Korc for technical support with molecular cloning; Jon Ivins for suggestions on optimizing AP detection methods; and Beverly Zak and Jon Ivins for thorough and helpful comments on this manuscript.

[3] R. L. Chen and A. D. Lander, unpublished observations.

REFERENCES

1. Kjellen, L., and Lindahl, U. (1991) *Annu. Rev. Biochem.* **60,** 443–475
2. Gallagher, J. T., Turnbull, J. E., and Lyon, M. (1992) *Int. J. Biochem.* **24,** 553–560
3. Prydz, K., and Dalen, K. T. (2000) *J. Cell Sci.* **113,** 193–205
4. Fritz, T. A., Gabb, M. M., Wei, G., and Esko, J. D. (1994) *J. Biol. Chem.* **269,** 28809–28814
5. Rohrmann, K., Niemann, R., and Buddecke, E. (1985) *Eur. J. Biochem.* **148,** 463–469
6. Inki, P., Gomez, M., Quintanilla, M., Cano, A., and Jalkanen, M. (1992) *Lab. Invest.* **67,** 225–233
7. Kokenyesi, R., and Bernfield, M. (1994) *J. Biol. Chem.* **269,** 12304–12309
8. Uhlin-Hansen, L., Kusche-Gullberg, M., Berg, E., Eriksson, I., and Kjellen, L. (1997) *J. Biol. Chem.* **272,** 3200–3206
9. Toyama-Sorimachi, N., Kitamura, F., Habuchi, H., Tobita, Y., Kimata, K., and Miyasaka, M. (1997) *J. Biol. Chem.* **272,** 26714–26719
10. Nadanaka, S., Kitagawa, H., and Sugahara, K. (1998) *J. Biol. Chem.* **273,** 33728–33734
11. Zhang, L., and Esko, J. D. (1994) *J. Biol. Chem.* **269,** 19295–19299
12. Zhang, L., David, G., and Esko, J. D. (1995) *J. Biol. Chem.* **270,** 27127–27135
13. Dolan, M., Horchar, T., Rigatti, B., and Hassell, J. R. (1997) *J. Biol. Chem.* **272,** 4316–4322
14. Litwack, E. D., Ivins, J. K., Kumbasar, A., Paine-Saunders, S., Stipp, C. S., and Lander, A. D. (1998) *Dev. Dyn.* **211,** 72–87
15. Lories, V., Cassiman, J. J., Van den Berghe, H., and David, G. (1992) *J. Biol. Chem.* **267,** 1116–1122
16. Shworak, N. W., Shirakawa, M., Mulligan, R. C., and Rosenberg, R. D. (1994) *J. Biol. Chem.* **269,** 21204–21214
17. Bernfield, M., Gotte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., and Zako, M. (1999) *Annu. Rev. Biochem.* **68,** 729–777
18. Cheng, H. J., Nakamoto, M., Bergemann, A. D., and Flanagan, J. G. (1995) *Cell* **82,** 371–381
19. Litwack, E. D., Stipp, C. S., Kumbasar, A., and Lander, A. D. (1994) *J. Neurosci.* **14,** 3713–3724
20. Herndon, M. E., and Lander, A. D. (1990) *Neuron* **4,** 949–961
21. Cheng, H. J., and Flanagan, J. G. (1994) *Cell* **79,** 157–168
22. Linker, A., and Hovingh, P. (1972) *Biochemistry* **11,** 563–568
23. Rapraeger, A., and Yeaman, C. (1989) *Anal. Biochem.* **179,** 361–365
24. Flanagan, J. G., and Leder, P. (1990) *Cell* **63,** 185–194
25. Tartaglia, L. A., Dembski, M., Weng, X., Deng, N., Culpepper, J., Devos, R., Richards, G. J., Campfield, L. A., Clark, F. T., and Deeds, J. (1995) *Cell* **83,** 1263–1271
26. Berger, J., Hauber, J., Hauber, R., Geiger, R., and Cullen, B. R. (1988) *Gene* (*Amst.*) **66,** 1–10
27. Ham, R. G. (1965) *Proc. Natl. Acad. Sci. U. S. A.* **53,** 288–293
28. Herndon, M. E., Stipp, C. S., and Lander, A. D. (1999) *Glycobiology* **9,** 143–155
29. Turnbull, J. E., and Gallagher, J. T. (1988) *Biochem. J.* **251,** 597–608
30. Mertens, G., Van der Schueren, B., van den Berghe, H., and David, G. (1996) *J. Cell Biol.* **132,** 487–497
31. Lander, A. D., Stipp, C. S., and Ivins, J. K. (1996) *Perspect. Dev. Neurobiol.* **3,** 347–358
32. Schwarz, K., Breuer, B., and Kresse, H. (1990) *J. Biol. Chem.* **265,** 22023–22028
33. Mann, D. M., Yamaguchi, Y., Bourdon, M. A., and Ruoslahti, E. (1990) *J. Biol. Chem.* **265,** 5317–5323
34. Brahimi-Horn, M. C., Deudon, E., Paul, A., Mergey, M., Mailleau, C., Basbaum, C., Dohrman, A., and Capeau, J. (1994) *Eur. J. Cell Biol.* **64,** 271–280
35. Saunders, S., Paine-Saunders, S., and Lander, A. D. (1997) *Dev. Biol.* **190,** 78–93
36. Cano-Gauci, D. F., Song, H. H., Yang, H., McKerlie, C., Choo, B., Shi, W., Pullano, R., Piscione, T. D., Grisaru, S., Soon, S., Sedlackova, L., Tanswell, A. K., Mak, T. W., Yeger, H., Lockwood, G. A., Rosenblum, N. D., and Filmus, J. (1999) *J. Cell Biol.* **146,** 255–264
37. Nakato, H., Futch, T. A., and Selleck, S. B. (1995) *Development* **121,** 3687–3702