REVIEW

Therapeutically targeting protein–glycan interactions

A Rek[1], E Krenn[2] and AJ Kungl[1,2]

[1]*ProtAffin Biotechnologie AG, Graz, Austria, and* [2]*Institute of Pharmaceutical Sciences, University of Graz, Graz, Austria*

Glycosylation is the most common form of post-translational modifications by which oligosaccharide side chains are covalently attached to specific residues of the core protein. Especially O-linked glycan structures like the glycosaminoglycans were found to contribute significantly to many (patho-)biological processes like inflammation, coagulation, cancer and viral infections. Glycans exert their function by interacting with proteins thereby changing the structure of the interacting proteins and consequently modulating their function. Given the complex nature of cell-surface and extracellular matrix glycan structures, this therapeutic site has been neglected for a long time, the only exception being the antithrombin III-glycan interaction which has been sucessfully targeted by unfractionated and low-molecular weight heparins for many decades. Due to the recent breakthrough in the '–ome' sciences, among them proteomics and glycomics, protein–glycan interactions became more amenable for therapeutic approaches so that novel inhibitors of this interaction are currently in preclinical and clinical studies. An overview of current approaches, their advantages and disadvantages, is given and the promising potential of pharmacologically interfering with protein–glycan interactions is highlighted here.
British Journal of Pharmacology (2009) **157,** 686–694; doi:10.1111/j.1476-5381.2009.00226.x; published online 9 April 2009

Structure–function considerations of glycans

Among post-translational modifications of eukaryotic proteins, glycosylation is the most common and the most diverse modification. It is therefore obvious, as each cell spends a tremedous biosynthetic effort on glycosylation, that these modifications are of general importance for many biological processes such as stability and folding assistance, embryonic development, antigenic and microbial recognition, inflammatory response and metastasis (Rudd *et al.*, 2001; Middleton *et al.*, 2002; Fears and Woods, 2006). The wide diversity of glycoforms found in nature, in addition to frequent dramatic consequences of altering these types of glycosylation – resulting in the loss of particular protein functions or in the degradation of the entire protein – has highlighted our limited understanding of structure–function relationships of complex glycans. Recently, tools to unravel glycan function on the basis of their saccharide building blocks have been developed as have methods for investigating the signals directing their biosynthetic enzymes (Varki, 1993; Prydz and Dalen, 2000; Taylor and Gallo, 2006). This opens ways for novel approaches to target this pharmacologically highly interesting site (Gesslbauer and Kungl, 2006).

Glycoproteins contain usually one or more oligosaccharide chains covalently attached to asparagine residues (termed N-linked) or to serine, threonine, hydroxylysine or hydroxyproline residues (termed O-linked). They are found ubiquitously in nature either as soluble (intracellular or extracellular) or as membrane bound molecules. There is a great structural variety based on the type, length and linkage of the carbohydrate components as well as on the degree of saturation of potential glycosylation sites on the protein itself (Berg *et al.*, 2002). A major characteristic of most glycoproteins is that the carbohydrate component constitutes a much smaller percentage of the overall molecular weight of the glycoprotein in relation to the core protein. This is in clear contrast to the so-called proteoglycans (PGs) in which the protein core is modified to a much higher degree by long polysaccharide chains commonly of the glycosaminoglycan (GAG) type (Mulloy and Rider, 2006).

Protein glycosylation is an enzymatic process which takes place inside the lumen of the endoplasmic reticulum (ER) and the Golgi apparatus. N-linked glycosylation is a co-translational event accompanying protein synthesis in ribosomes located in the ER, whereas O-linked glycosylation is a post-translational process characterized by the stepwise addition of sugar residues directly to the protein. The most common type of O-linked glycans contain an initial

Targeting protein–glycan interactions
A Rek *et al*
687

N-acetylgalactosamine residue bound to the Ser/Thr residues (=mucine-type glycans). Other O-linked glycans are known to include glucosamine, xylose, galactose, fucose or mannose as the primary sugar. Furthermore, O-linked glycoproteins are very often large proteins (>200 kDa) with clusters of glycosylation chains which are comparatively less branching than N-glycans. Chain elongation and termination, if performed, are carried out by various glycosyltransferases localized in different regions of the Golgi. Termination of O-linked glycans usually includes Gal, GlcNAc, GalNAc, Fuc or sialic acid (Hang and Bertozzi, 2005; Lehle *et al.*, 2006).

N-linked glycans are found among lysosomal enzymes, cytokines, hormones, growth factors and their receptors, plasma proteins, matrix proteins, membrane proteins and receptors, etc. (McKillop *et al.*, 2000; Duchesne *et al.*, 2006; Van Acker *et al.*, 2006). O-Linked glycans are prevalent in most secretory cells and tissues. They are present in high concentrations in the zona pelucida surrounding mammalian eggs and may funtion as sperm receptors (ZP3 glycoprotein) (Shur *et al.*, 2006). O-linked glycans are also involved in hematopoiesis, inflammation response mechanisms and the formation of ABO blood antigens (Rydberg, 2001). Nonelongated O-GlcNAc groups have been recently shown to be related to phosphorylation states and dynamic processing related to cell signalling events in the cell (Whelan and Hart, 2003). From a pharmacological and therapeutic point of view, O-linked glycans, and here especially the GAGs, have a much higher potential for targeted intervention. The reason for this being the molecular interaction framework in which GAGs are involved and which ultimately modulate the function of a large number of signalling molecules and extracellular enzymes (see below).

The structural distribution of glycans is strongly influenced by the tissue-specific expression of the biosynthetic enzymes, their activity and specificity as well as of sugar nucleotide pool composition (Dennis *et al.*, 1999). Glycan patterns therefore reflect a certain biological/pathological state of a cell or a tissue. We have recently shown that human microvascular endothelial cells change their GAG sulfation pattern follwing an inflammatory (TNF-alpha) trigger (Krenn *et al.*, 2008). As the methods for analysing cell-/tissue-derived glycans are still very limited, we have applied so-called glyco-genomic methods to infer the glycan expression pattern of endothelial cells from the expression pattern of GAG biosynthetic enzymes as determined by qPCR analysis.

Biosynthesis of O-linked glycans: heparan sulphate (HS) GAGs as a pharmacologically interesting case study

The protein core components of PGs are synthesized in ribosomes to be then translocated to the rough ER where a xylosyltransferase initiates the synthesis of the linker tetrasacharide by adding a xylose to a serine residue of the protein core. Two galactose residues are subsequently added in the cis or medial Golgi to the Xyl by galactosyltransferase I and galactosyltransferase II. The fourth residue, completing the linker tetrassacharide, is a GlcA added by glucuronyltrans-

ferase I and occurs in the trans-Golgi, the final location for all subsequent reactions. The addition of the fifth saccharide determines whether the GAG chain becomes chondroitin sulfate (CS)/DS or HS/heparin. GAG type, length of the chain(s), conformational flexibility and particularly the specific GAG sequence/structure determine the biological function of the glycan part of the PG. The unique structural design, which in turn determines the specific binding properties of the PG, is generated during biosynthesis by the concerted action of a complex set of enzymes (Esko and Selleck, 2002; Gesslbauer *et al.*, 2007) which is supposed to be non-template dependent and therefore not related to a so-called glyco-code (Habuchi *et al.*, 2004). During chain elongation, the nascent GAG chain is modified by an epimerase, converting GlcA into IdoA, and several sulfotransferases adding sulfate groups to distinct positions. Chain elongation and modification require an array of distinct enzymes for the HS and the CS pathway (see Figure 1). This conventional view of GAG biosynthesis, which is based on sequential modifications of the chain, has been recently challenged by the postulation of a so-called GAGososme (Ledin *et al.*, 2006) in which the biosynthetic enzymes are topologically clustered within the Golgi membranes and can therefore act in a more concerted and regulated way. The mature HS chain can also be edited by the action of endosulfatases and heparanase. Especially, the enzymes involved in the generation of the sulfation pattern exist in several isoforms with divergent activities, substrate specificities and tissue distribution. Modulation in GAG structure is therefore likely to be achieved, at least to some extend, by the differential regulation of expression of a certain repertoire of modifying enzymes. In Figure 1, the overall scheme of HS synthesis is displayed indicating the shaping of AT-III and FGF2 binding sites.

Specific shaping of GAG structures was found to occur during pathogenesis, during development or in response to extracellular signals such as growth factors (Carter *et al.*, 2003). Among the various physiological processes (like cell proliferation, differentiation, morphogenesis, angiogenesis, etc.) in which HS is involved, the interaction with extracellular signalling molecules deserves special attention. Among these signalling molecules are growth factors, and chemokines, lipid- and membrane-binding proteins, adhesions proteins, and certain proteases and esterases, as well as HS degrading enzymes (Capila and Linhardt, 2002; Handel *et al.*, 2005; Vlodavsky *et al.*, 2007). HS plays an essential role in the regulation of inflammatory processes, cell growth and differentiation, lipid transport and clearance, cell–cell and cell–matrix interactions and blood coagulation processes (Kolset and Salmivirta, 1999; Knox and Whitelock, 2006). The association of chemokines with HS helps to stabilize concentration gradients across the endothelial surface thereby providing directional clues for migrating leukocytes in inflammation (e.g. atherosclerosis, arthritis). HS was also shown to protect chemokines from proteolytic degradation and to induce their oligomerization, thus promoting local high concentrations in the vicinity of their G-coupled signalling receptors (Johnson *et al.*, 2005). FGF1 and FGF2 were the first members of the fibroblast growth factor family to be discovered and their interaction with HS was intensively investigated. It has been established that HS is required for
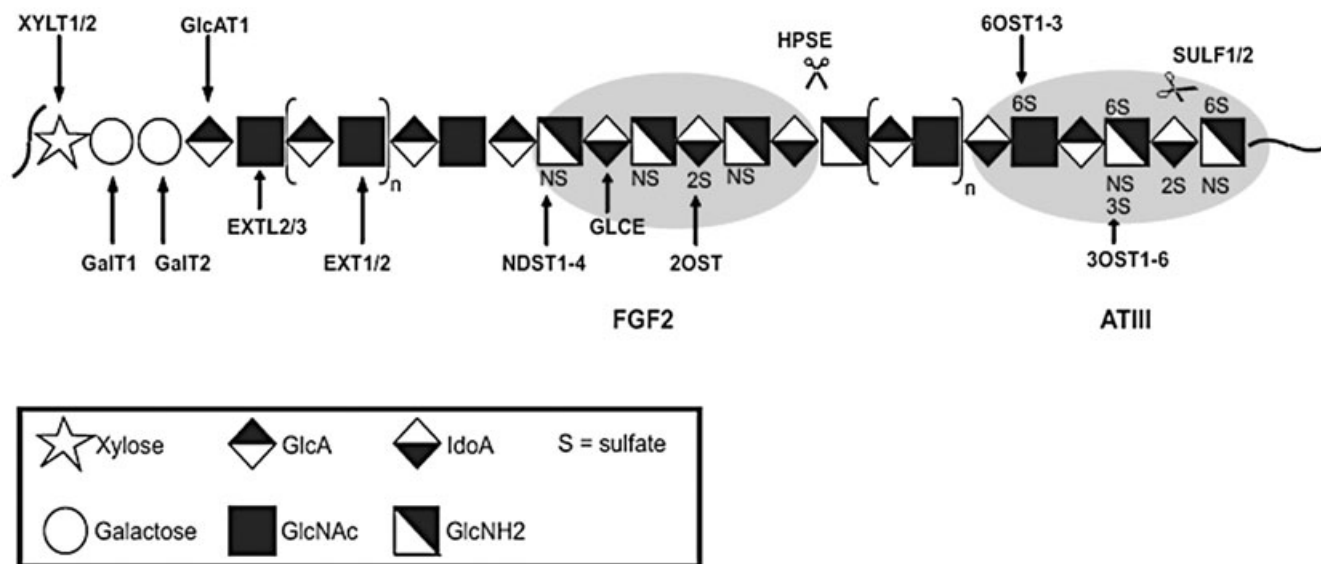
Figure 1 Scheme of HS-biosynthesis. Displayed is the consecutive action of the various biosynthetic enzymes durin chain initiation, polymerization and modification (the recently proposed model of a so-called GAGosome, in which the biosynthetic steps occur in a parallelized way (Ledin *et al.*, 2006), has not been considered in this figure). Enzymes involved in HS biosynthesis are: XYLT1, xylosyltransferase-1; XYLT2, xylosyltransferase-2; GalT1, galactosyltransferase-1; GalT2, galactosyltransferase-2; GlcAT1, glucuronosyltransferase-1; EXT1, exostosin-1; EXT2, exostosin-2; EXTL2, exostosin-like-2; EXTL3, exostosin-like-3; NDST1-4, N-deacetylase/N-sulfotransferase-1, -2, -3, -4; GLCE, D-glucuronyl C5-epimerase; 2OST, 2-O-sulfotransferase; 3OST1-6, 3-O-sulfotransferase-1, -2, -3A, -3B, -4, -5, -6; 6OST1-3, 6-O-sulfotransferase-1, -2, -3; SULF1, extracellular sulfatase Sulf-1; SULF2, extracellular sulfatase Sulf-2; HPSE, heparanase. Two typical results of taylored HS biosynthesis are given in the form of the FGF2- and the ATIII-specific binding sites. HS, heparan sulphate.

FGF to effectively activate its receptor and several models have been described trying to explain FGF/HS/FGF-receptor complex formation and its stoichiometry. For FGF2, to form the ternary signalling complex, there must be FGF-2, FGF receptor, and an HS chain having appropriately spaced 2-O and 6-O sulfation. FGF-2 released from HS chains in the extracellular matrix (ECM) by cleavage through mammalian heparanase has been shown to promote angiogenesis for tumour growth (Sanderson *et al.*, 2005; Harmer, 2006).

Localization of O-linked glycans: they are part of PGs

Proteoglycans are a special class of glycanated proteins which consist of a core protein with one or more covalently attached GAG side chains. Their carbohydrate structure clearly distinguishes them from 'ordinary' glycoproteins: repeating disaccharide units composed of an hexosamine and an uronic acid are arranged to form the long linear GAG polysaccharide chains. Depending on the type of the disaccharide building blocks, the PG superfamily is subdivided into three major groups: (i) chondroitin/dermatan sulfate proteoglycans (CSPGs); (ii) heparin/heparin sulfate proteoglycans (HSPGs); and (iii) keratan sulfate proteoglycans (KSPGs) (Mulloy and Rider, 2006; Bishop *et al.*, 2007). All three GAG structures are strongly negatively charged due to their high content of sulfate groups and uronic acids. Although there are intracel-

lular types, most PGs are directed to the cell surface as well as to the ECM including cartilage, brain, skin, connective tissue, basement membranes and blood vessels. The biological functions of PGs are diverse ranging from purely structural function in the ECM, mechanical tissue function, influence in metastasis and their interaction with a variety of extracellular ligands, such as growth factors, chemokines and adhesions molecules (Vlodavsky *et al.*, 2002; Kolset *et al.*, 2004; Cattaruzza and Perris, 2005; Celie *et al.*, 2005; Proudfoot, 2006). The common motif among the various functions is the interaction of the PG, in most cases only the GAG part, with other proteins (see below). The subsequent structural rearrangement of the interacting protein is usually the trigger for a specific biological function.

Cell surface HSPGs are subdivided into the glypican and the syndecan families, which are the major representatives of this class, betaglycan and CD44 being minor representatives. Glypicans are glycosylphosphatidylinositol (GPI)-linked molecules containing exclusively HS GAGs. The GPI anchor is responsible for a different turnover and involvement in different metabolic pathways compared with the integral PGs. It also affects their localization and distribution of the glypicans. There are six known members of the glypican family in mammals: glypican-1, glypican-2 (cerebroglycan), glypican-3 (OCI-5), glypican-4 (K-glypican), glypican-5 (GCP5) and glypican-6 (David, 1993; Fransson, 2003). The syndecan family contains four members: syndecan-1, syndecan-2 (fibroglycan), syndecan-3 (N-syndecan) and syndecan-4

(amphiglycan or ryudocan) which have been described in vertebrates. They are type I glycoproteins with a single membrane spanning protein core comprising a large ectodomain and a highly conserved transmembrane domain. Their short cytoplasmatic parts contain two conserved domains interrupted by a variable region, which is specific for each syndecan family member, thus allowing both common and core-specific interactions with various cytoplasmic molecules. Syndecans usually carry three HS chains close to their N-terminus, although substitutions with CS are possible (Gotte, 2003; Tkachenko *et al.*, 2005).

Proteoglycans modified by CS chains include aggrecan, neurocan, brevican, bamacan, a CD44 isoform, syndecans, betaglycan and serglycin (Taylor and Gallo, 2006). DS is mainly found in skin and PGs bearing DS chains are versican, decorin, biglycan and endocan (Trowbridge and Gallo, 2002). Many of the PGs bearing CS chains are additionally modified by GAG chains of either HS or KS, for example, syndecans or aggrecan respectively. Decorin, however, a secretory product in the ECM, is one example of an exclusively CS/DS PG. Aggrecan has a molecular mass of about 2500 kDa and is a large KS/CSPG in cartilagenous tissue. Its distribution pattern is relatively restricted to a few tissues including brain, aorta and tendon, in addition to cartilage. The core protein of 210–250 kDa binds hyaluronic acid and forms a supramolecular complex together with the linker protein. Aggrecan provides a strongly hydrated space filling gel due to the large number (more than 100) of polyanionic GAG chains covalently attached to the protein core and thereby contributes to the physical strength of cartilagenous tissue (Roughley, 2006).

Pharmacological driving force: protein–GAG interactions and the glycan-protein network (GPN)

While PGs are involved in a plethora of physiological processes, much has been speculated about the individual roles of the core protein and the attached GAG chains. Although we have recently shown that the chemokine CXCL8 binds to both, the GAG sidechain and the core protein of syndecan-2 (Halden *et al.*, 2004), and is therefore able to induce downstream signalling via its syndecan co-receptor into endothelial cells (S.F. Falsone *et al.*, manuscript submitted), it is generally accepted that mainly the glycan side chains on PGs are responsible for the biological functions of PGs. This is achieved by immobilizing and by structurally influencing proteins which bind to these complex carbohydrate chains. Interestingly, although the affinity of proteins towards GAG ligands is commonly rather low, that is, in the μmol·L^{-1} range, a significant structural change can be induced in the protein (Goger *et al.*, 2002). The function of the interacting protein is thereby modulated either by inducing a direct conformational change (Goger *et al.*, 2002), sometimes by the induction of homo- or hetero-oligomerization (Hoogewerf *et al.*, 1997; Weber and Koenen, 2006). Of special pharmacological interest in the context of protein–GAG interactions is the cell surface of endothelial cells. These cells, being the coating of the (micro)vasculature, are able to secrete, release or express
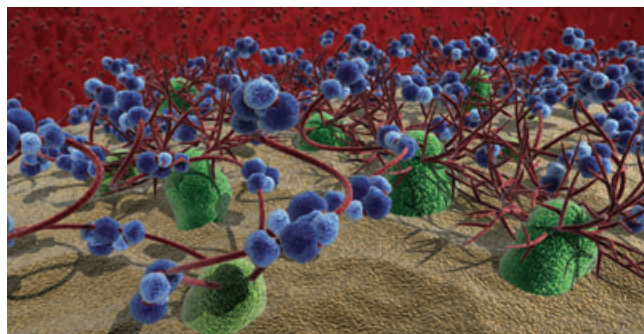

Figure 2 Cell surface of endothelial cells. In green, (trans-)membrane proteins are shown; branched and short red tubes represent N-linked glycans, long and unbranched red tubes represent O-linked (glycoasmino-)gycans; in blue, glycan-binding proteins are shown, different shades refer to different proteins and thus to hetero- or homo-oligomerization.

growth factors and chemokines, adhesions proteins, proteases and esterases, as well as HS degrading enzymes into the serum which bind to HS or CS, the proto-typic and dominating glycan chains on most eukaryotic cells, thereby attracting and activating distant cells like leukocytes or metastatic cells or releasing cells in the course of metastasis or angiogenesis. Consequently, the surface of endothelial cells is characterized not only by the various membrane proteins and their short and branched N-linked glycan chains, the glycocalyx of a cell is rather constituted by a wealth of charged glycan (GAG) chains which, in addition, bind a multitude of proteins, thereby creating a dense matrix of membrane proteins, glycan chains and immobilized soluble proteins (see Figure 2). We define this matrix as the GPN which ultimately defines the active and passive signalling capacity of a given cell, as it relates secreted and presented proteins to the functional environment of the cell.

If the GPN is a signalling fingerprint of a cell, then the protein–glycan interactions are proposed to be of significant specificity. In their seminal work, Cardin and Weintraub (1989) have outlined two protein consensus sequences which are responsible for GAG binding. Their report describes defined arrays of basic amino acid clusters (lysine, arginine and histidine) characterized by the consensus sequences, XBBXBX and XBBBXXBX, 'where B is the probability of a basic residue and X is a hydropathic residue'. More recently, Hileman *et al.* (Hileman *et al.*, 1989) have extended this list by the new consensus sequence TXXBXXTBXXXTBB, where turns (T) are suggested to bring basic interacting amino acid residues into proximity. But what about protein-specific GAG sequences? Less is known about specific GAG sequences mainly because (bio)chemical methods for preparing and amplifying biologically relevant (i.e. cell- or tissue-derived) and protein-related GAGs are lacking, although highly sensitive technologies for sequencing these carbohydrates have been introduced by the Gallagher group in the late 1990s (Merry *et al.*, 1999; Turnbull *et al.*, 1999). However, no high throughput sequencing methods like for mass spectrometry (MS)/MS-based protein sequencing in proteomics have yet been established (see below). Table 1 shows specific GAG sequences which were to date found for proteins of various

Table 1 Protein-specific glycosaminoglycan sequences

GAG-sequence	Interacting protein	References
GlcNAc/NS(6S)-GlcA-GlcNS(3,6S)-IdoA(2S)-GlcNS(6S)	Antithrombin III	Oscarsson LG et al. (1989). J Biol Chem **264** (1): 296–304.
UA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)	Annexin V	Capila I et al. (2001). Structure **9**: 57–64.
UA-GlcNS-IdoA(2S)-GlcNAc-UA(2S)-GlcNS-IdoA(2S)-GlcNH$_2$(3,6S)	herpes simplex 1 envelope glycoprotein D (gD)	Liu J et al. (2002). J Biol Chem **277** (36): 33456–33467.
IdoA-GlcNAc(6S)-GlcA-GlcNS(3,6-S)-IdoA(2S)-IdoA2S-GlcNS6S-	Heparanase	Gong F et al. (2003). J Biol Chem **278** (37): 35152–35158.
	HSulf1, HSulf2	Morimoto-Tomita et al. (2002). J Biol Chem **277** (51): 49175–49185.
GlcNS-IdoA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS	VEGF 165	Robinson CJ et al. (2006). J Biol Chem **281** (3): 1731–1740.
GlcA-GlcNS-[IdoA(2S)-GlcNS]$_5$-IdoA-GlcNAc	FGF2	Turnbull et al. (1992). J Biol Chem **267** (15): 10337–10341.
-IdoA(2S)-GlcNS(6S)-IdoA(2S)-	FGF1	Kreuger J et al. (2001). J Biol Chem **276** (33): 30744–30752.
UA-GlcNS-[UA(2S)-GlcNS(6S)]$_3$	FGF4, FGF7	
UA-GlcNS-[UA-GlcNS(6S)]$_3$	FGF10	
UA-GlcNS-[UA(2S)-GlcNS]$_3$, UA-GlcNS-[UA-GlcNS(6S)]$_3$	FGF18, HGF	Ashikari S et al. (2004). J Biol Chem **279** (13): 12346–12354.
UA-GlcNS/NAc-IdoA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)-UA-GlcNS/NAc-	HGF	Ashikari S et al. (1995). J Biol Chem **270** (49): 29586–29593.
UA-GlcNS/NAc-UA-GlcNS/NAc-IdoA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)-		
UA-GlcNS/NAc-IdoA(2S)-GlcNS(6S)-UA-GlcNS/NAc-IdoA(2S)-GlcNS(6S)-		
-GlcA-GlcNAc/S-[IdoA(2S)-GlcNS(6S)]$_3$-[GlcA-GlcNAc/S]$_{6-7}$-[IdoA(2S)-GlcNS(6S)]$_3$-GlcA-GlcNAc/S	IL8	Spillmann et al. (1998). J Biol Chem **273** (25): 15487–15493.
(UA-GlcN)$_8$ containing 7S and 2Ac	MCP1, MCP2, MCP3	Schenauer MR et al. (2007). J Biol Chem **282** (35): 25182–25188.
IdoA(2S)-GlcNH$_2$(3S ± 6S) within dp8	CyPB	Vanpouille C et al. (2007). J Biol Chem **282** (33): 24416–24429.

GAG, glycosaminoglycan.

classes demonstrating the general principle of the 'specificity' within protein–GAG interactions. However, based on more general findings (Kreuger et al., 2006) but also based on findings in our own lab, we tend to use 'selectivity' rather than specificity for describing this interaction, which allows for a certain degree of degeneracy during recognition of proteins and glycan molecules.

Identification of protein-specific GAG sequences

The pioneering work by Lindahl and co-workers has led to the identification of the first protein-specific GAG sequence, namely the heparin pentasaccharide structure required for antithrombin-III binding (Lindahl et al., 1989). Since then, only few more protein-specific GAG oligosaccharides have been added to this list (see Table 1). Traditionally, identification of protein-specific GAG oligosaccharides was accomplished in a ligand-biased manner, that is, by screening a naturally derived, diversified and size-defined GAG oligosaccharide library with respect to target protein binding. This approach, due to the limited size of the oligosaccharide library, is incomprehensive and therefore the very specific GAG ligand for a given protein may not (or not in sufficient amounts for detection) be contained in the screened library. The 'classical' methods used to obtain and to characterize protein-speficic GAG oligos include gel electrophoresis, affinity- and size-exclusion chromatography, filter binding and competition assays, microcalorimetry, and surface plasmon resonance. Recent developments in liquid chromatography and capillary electrophoresis coupled to MS in combination with subtle bioinformatic tools allow nowadays a more unbiased and efficient identification of protein-specific GAG oligos (Zamfir et al., 2004; Saad et al., 2005; Yu et al.,

2005; Yu et al., 2006). Using tandem MS, the group of Leary was able to unravel the structure of a number of HS oligos and to identify oligos which bind to specific chemokines (Yu et al., 2005). This approach is still limited by origin and natural diversity of the GAG oligosaccharide library used to select for protein-specific ligands. However, because of the general sensitivity of MS methods, a number of oligosaccharide libraries can be screened with a much higher probability to pick the specific GAG ligand even if it occurs with low abundancy. This means the dawning of the glycomic age also for glycan analyses. In the group of Sasisekharan, an MS-based glycan fingerprinting method combined with bioinformatical integration of data sets – a method using the so-called property-encoded nomenclature which was developed to handle many data sets and to extract information on the investigated saccharide structure – was successfully applied to various linear glycans including the AT-III-specific heparin epitope (Shriver et al., 2000). Enzyme protection assays were shown to be especially useful for the identification of GAG ligands of dimeric and multimeric proteins such as PF4 and MIP-1α (Stringer and Gallagher, 1997; Stringer et al., 2002). By applying heparinase and/or chondroitinase to protein–ligand complexes, also distant protein binding sites on the GAG ligand have been identified like the split S-domains for PF4 and long sulfated regions for MIP-1α. A similar approach has been used to identify the HS motif for interferon-γ (Lortat-Jacob et al., 1995).

Therapeutic approaches

From Table 1, it becomes obvious that a multitude of pharmacologically highly interesting target proteins need a selective GAG partner to exert their physiological function which

Targeting protein–glycan interactions
A Rek *et al*
691

are therefore potential therapeutic targets (Gesslbauer and Kungl, 2006; Lindahl, 2007). In general, as outlined above, protein–glycan interactions are considered to occur with low affinity, that is, in the low micromolar range, although higher affinities (in the nanomolar range) especially for growth factors have been observed (Mohammadi and Olsen, 2005). There are consequently various ways to interfere with protein–glycan interactions (see Figure 3). The 'classical' approach to this problem is to chemically and/or enzymatically synthesize the target protein-specific glycan structure (see Figure 3B) or a mimetic thereof and to test for and to optimize inhibitory activity. Despite the initial success of Sanofi and Organon in co-developing various forms of the AT-III-specific heparin pentasaccharide as a very efficient anti-thromboticum, resulting in Fondaparinux (Arixtra®), an a priori GAG synthesizing approach is currently less pursued by large pharmaceutical companies but rather by biotech companies like Momenta Pharmaceuticals Inc. and Endotis Pharma. Momenta's M118 is a rationally engineered heparin which is modified and fractionated so to provide only sites with anti-coagulant activity. Endotis' EP37 was synthesized, using the company's unique oligosaccharide synthesis platform, for oral prevention of venous thromboembolism and stroke prevention in patients with atrial fibrillation. Upcoming on the glycomics horizon are miniaturized on-chip synthesis technologies which use combinatorial or parallel chemistry to generate a multitude of glycan oligosaccharides for biological activity screening (de Paz *et al.*, 2007). These methods, however valuable they are for performing medium to highthroughput first screens and to find suitable hits, do not solve the problem of finally synthesizing the hit oligosaccharides in sufficient amounts needed for clinical trials or to feed the market for such drugs.

The typical biotech approach, that is, raising monoclonal antibodies against the glycan binding target protein, seems not to be generally applicable for glycan-binding proteins. As glycan-binding proteins are supposed to act only (or only in the desired fashion) when bound to the glycan, any mAb raised against the soluble protein might fail to inactivate the glycan-bound protein simply because the structural change is prone to induce epitope loss and/or epitope masking due to the high charge of the GAG ligand (see Figure 3A.1). This was most probably the reason for the failure of the anti-IL-8 antibody ABX-IL8 from Abgenix Inc. in clinical phase II, as the authors stated themselves that ABX-IL8 was not able to bind IL-8 localized on endothelial cells, that is, bound to GAGs (Yang *et al.*, 1999). On the other hand, it proved to be very difficult to raise monoclonal antibodies against protein-specific GAG sequences. Although the group of van Kuppevelt was successful in generating antibodies against various tissue-
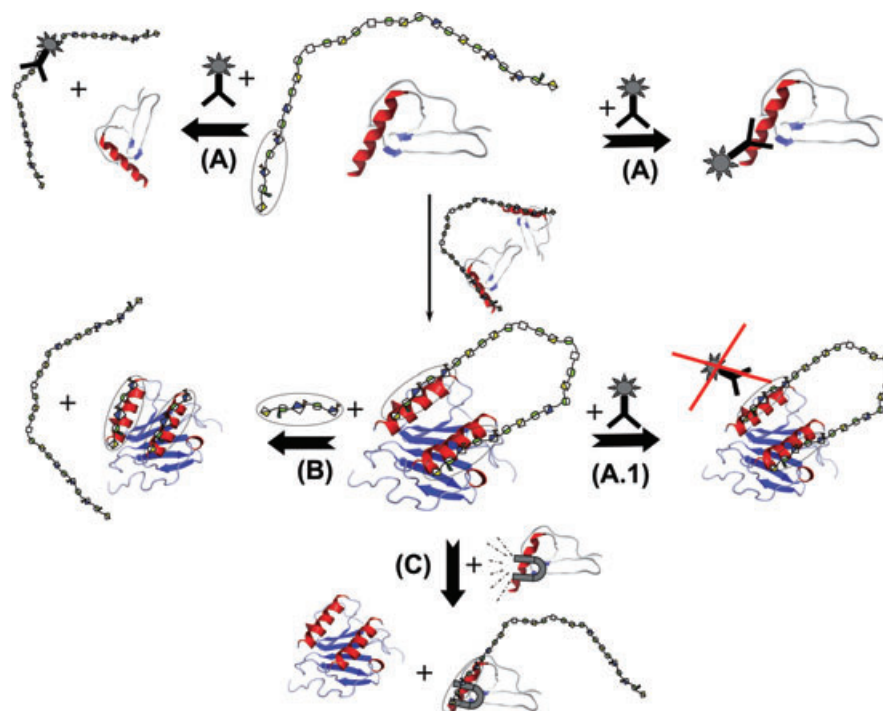

Figure 3 Different ways for therapeutic intervention in protein–GAG signalling. ⬙ sequence-specific GAG – Oligo within a proteoglycan side chain; ⫠ antibody. (A) Antibody approach. The binding of a protein to a (specific) GAG chain can be prevented by an antibody to either the (proteo-)glycan or the protein interaction partner. (A.1) The GAG-bound protein may not be targetable by the antibody raised against the soluble protein as structural rearrangements of the protein upon GAG binding as well as the change of overall/surface charge influence/destroy the antibody binding epitope. (B) Glycan- or glycan mimetic-based approach. The protein–GAG interaction can be antagonized by the addition of the synthesized protein-specific GAG sequence (or a mimetic thereof) which displaces the target protein from the natural GAG ligand. (C) Protein-based approach. The protein–GAG interaction can be antagonized by the addition of a mutant form of the target protein which exhibits higher affinity towards the natural GAG ligand and thereby displaces the wild-type target protein from the GAG chain. GAG, glycosaminoglycan.

specific GAGs by phage-display (Lensen et al., 2005), no thera-peutic antibody is yet available which inhibits specific protein–GAG interaction. This could be attributed to the low affinity of these phage display-derived antibodies and/or to the still low selectivity of these molecules. Antibodies against entire PGs have been available for quite some time but their potential for treatment of glycan-related diseases is not yet clearly foreseeable (Fjeldstad and Kolset, 2005). As these anti-bodies were raised either against only the recombinant core protein of PGs or against the entire molecule (core protein plus glycan chains), the effect on disrupting protein–glycan interactions remains elusive at this point. Another interesting therapeutic approach addressing protein–GAG interactions by knocking out the GAG interaction sites is the various chemokine mutants developed by Serono (Proudfoot et al., 2008; Shahrara et al., 2008). The group of Amanda Proudfoot engineered several chemokines towards lower or no GAG binding affinity with the aim to block the chemokines' GPCR receptors on neutrophils with these GAG knockout mutants, which consequently prevents leukocyte activation and firm adhesion by GAG-bound wild-type chemokines.

We have developed a protein-based technology to create a novel class of glycan-directed therapeutics which circumvents the above outlined problems for antagonizing protein/GAG interactions (Potzinger et al., 2006). Our CellJammer® (Prot-Affin, Graz, Austria) approach takes advantage of the intrinsic specificity or selectivity (see above) of a GAG-binding protein for its ligand to create a protein-based GAG antagonist by engineering the protein towards higher GAG binding affinity (see Figure 3C). This is achieved by engineering basic amino acids into the GAG binding site of the selected protein in a way that the overall structure of the protein is conserved. This kind of affinity maturation is based on the fact that electro-static interactions are mainly responsible for (high) affinity whereas hydrogen bonds and van der Vaals contacts deter-mine specificity (as in the case of protein–DNA interactions). By this means, the precise nature of the glycan ligand does not need to be known as long as specific/selective GAG ligand recognition is conserved within the mutant protein. This can usually be verified by (i) competition experiments of mutant protein versus wild-type protein versus other (unspecific) GAG-binding proteins; and (ii) by comparative binding experiments of mutant versus wild-type protein using a (limited) library of synthetic and natural, fractionated GAG oligosaccharides. As a first proof of concept, we have devel-oped an CXCL8-based therapeutic, PA401, which exhibits a 100-fold higher affinity towards HS while keeping selective GAG ligand recognition compared with the wild-type chemokine. PA401 was found to bind also to CS but the differences in selectivity with respect to wtCXCL8 need to be explored in more detail in the future. This mutant protein was shown to efficiently displace wtCXCL8 from HS chains on endothelial cells, whereas other chemokines are displaced only at much higher concentrations of PA401. To this engi-neered 'dominant' (higher GAG binding affinity) characteris-tics of the chemokine, we have added a 'negative' characteristic in order to create a potent anti-inflammatory drug, otherwise a super-agonist of CXCL8 had been the result. For this purpose, the leukocyte attraction and activation domain of the chemokine have been knocked out by deleting six amino acids of the N-terminus. The dominant-negative CXCL8 mutant PA401 exhibited excellent inhibitory activity in a number of inflammatory animal models (A. Rek et al., manuscript in preparation) and is, among further proteins modified by the CellJammer® (ProtAffin) technology, cur-rently in preclinical development. Our approach represents the first protein-based GAG antagonist which specifically inhibits information flow from the HS-presenting endothe-lium to the attracted leukocyte.

Acknowledgements

Conflict of interest

A Rek and AJ Kungl are employees of ProtAffin Biotechnologie AG, Austria. The company has interest in the development of protein-based GAG antagonists.

References

Ashikari S, Habuchi H, Kimata K (1995). Characterization of heparan sulfate oligosaccharides that bind to hepatocyte growth factor. J Biol Chem 270: 29586–29593.

Ashikari-Hada S, Habuchi H, Kariya Y, Itoh N, Reddi AH, Kimata K (2004). Characterization of growth factor-binding structures in heparin/heparan sulfate using an octasaccharide library. J Biol Chem 279: 12346–12354.

Berg JM, Tymodczko JL, Stryer L (2002). Biochemistry. W. H. Freeman and Company: New York.

Bishop JR, Schuksz M, Esko JD (2007). Heparan sulphate proteogly-cans fine-tune mammalian physiology. Nature 446: 1030–1037.

Capila I, Linhardt RJ (2002). Heparin-protein interactions. Angew Chem Int Ed Engl 41: 391–412.

Capila I, Hernáiz MJ, Mo YD, Mealy TR, Campos B, Dedman JR et al. (2001). Annexin V–heparin oligosaccharide complex suggests heparan sulfate–mediated assembly on cell surfaces. Structure 9: 57–64.

Cardin AD, Weintraub HJ (1989). Molecular modeling of protein-glycosaminoglycan interactions. Arteriosclerosis 9: 21–32.

Carter NM, Ali S, Kirby JA (2003). Endothelial inflammation: the role of differential expression of N-deacetylase/N-sulphotransferase enzymes in alteration of the immunological properties of heparan sulphate. J Cell Sci 116: 3591–3600.

Cattaruzza S, Perris R (2005). Proteoglycan control of cell movement during wound healing and cancer spreading. Matrix Biol 24: 400–417.

Celie JW, Keuning ED, Beelen RH, Drager AM, Zweegman S, Kessler FL et al. (2005). Identification of L-selectin binding heparan sulfates attached to collagen type XVIII. J Biol Chem 280: 26965–26973.

David G (1993). Integral membrane heparan sulfate proteoglycans. FASEB J 7: 1023–1030.

Dennis JW, Granovsky M, Warren CE (1999). Protein glycosylation in development and disease. Bioessays 21: 412–421.

Duchesne L, Tissot B, Rudd TR, Dell A, Fernig DG (2006). N-glycosylation of fibroblast growth factor receptor 1 regulates

ligand and heparan sulfate co-receptor binding. *J Biol Chem* **281**: 27178–27189.

Esko JD, Selleck SB (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem* **71**: 435–471.

Fears CY, Woods A (2006). The role of syndecans in disease and wound healing. *Matrix Biol* **25**: 443–456.

Fjeldstad K, Kolset SO (2005). Decreasing the metastatic potential in cancers – targeting the heparan sulfate proteoglycans. *Curr Drug Targets* **6**: 665–682.

Fransson LA (2003). Glypicans. *Int J Biochem Cell Biol* **35**: 125–129.

Gesslbauer B, Kungl AJ (2006). Glycomic approaches toward drug development: therapeutically exploring the glycosaminoglyca-nome. *Curr Opin Mol Ther* **8**: 521–528.

Gesslbauer B, Rek A, Falsone F, Rajkovic E, Kungl AJ (2007). Proteogly-canomics tools to unravel the biological function of glycosamino-glycans. *Proteomics* **7**: 2870–2880.

Goger B, Halden Y, Rek A, Moesl R, Pye D, Gallagher J *et al*. (2002). Different affinities of glycosaminoglycan oligosaccharides for monomeric and dimeric interleukin-8: a model for chemokine regulation at inflammatory sites. *Biochemistry* **41**: 1640–1646.

Gong F, Jemth P, Escobar Galvis ML, Vlodavsky I, Horner A, Lindahl U *et al*. (2003). Processing of macromolecular heparin by hepara-nase. *J Biol Chem* **278**: 35152–35158.

Gotte M (2003). Syndecans in inflammation. *FASEB J* **17**: 575–591.

Habuchi H, Habuchi O, Kimata K (2004). Sulfation pattern in gly-cosaminoglycan: does it have a code? *Glycoconj J* **21**: 47–52.

Halden Y, Rek A, Atzenhofer W, Szilak L, Wabnig A, Kungl AJ (2004). Interleukin-8 binds to syndecan-2 on human endothelial cells. *Biochem J* **377**: 533–538.

Handel TM, Johnson Z, Crown SE, Lau EK, Proudfoot AE (2005). Regulation of protein function by glycosaminoglycans – as exemplified by chemokines. *Annu Rev Biochem* **74**: 385–410.

Hang HC, Bertozzi CR (2005). The chemistry and biology of mucin-type O-linked glycosylation. *Bioorg Med Chem* **13**: 5021–5034.

Harmer NJ (2006). Insights into the role of heparan sulphate in fibro-blast growth factor signalling. *Biochem Soc Trans* **34**: 442–445.

Hileman RE, Fromm JR, Weiler JM, Linhardt RJ (1989). Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *Bioessays* **20**: 156–167.

Hoogewerf AJ, Kuschert GSV, Proudfoot AEI, Borlat F, Clark-Lewis I, Power CA *et al*. (1997). Glycosaminoglycans mediate cell surface oligomerization of chemokines. *Biochemistry* **36**: 13570–13578.

Johnson Z, Proudfoot AE, Handel TM (2005). Interaction of chemok-ines and glycosaminoglycans: a new twist in the regulation of chemokine function with opportunities for therapeutic interven-tion. *Cytokine Growth Factor Rev* **16**: 625–636.

Knox SM, Whitelock JM (2006). Perlecan: how does one molecule do so many things? *Cell Mol Life Sci* **63**: 2435–2445.

Kolset SO, Salmivirta M (1999). Cell surface heparan sulfate proteogly-cans and lipoprotein metabolism. *Cell Mol Life Sci* **56**: 857–870.

Kolset SO, Prydz K, Pejler G (2004). Intracellular proteoglycans. *Biochem J* **379**: 217–227.

Krenn EC, Wille I, Gesslbauer B, Poteser M, van Kuppevelt TH, Kungl AJ (2008). Glycanogenomics: a qPCR-approach to investigate biological glycan function. *Biochem Biophys Res Commun* **375**: 297–302.

Kreuger J, Spillmann D, Li JP, Lindahl U (2006). Interactions between heparan sulfate and proteins: the concept of specificity. *J Cell Biol* **174**: 323–327.

Kreuger J, Salmivirta M, Sturiale L, Giménez-Gallego G, Lindahl U (2001). Sequence analysis of heparan sulfate epitopes with graded affinities for fibroblast growth factors 1 and 2. *J Biol Chem* **276**: 30744–30752.

Ledin J, Ringvall M, Thuveson M, Eriksson I, Wileén M, Kusche-Gullberg M *et al*. (2006). Enzymatically active N-deacetylase/N-sulfotransferase-2 is present in liver but does not contribute to heparan sulfate N-sulfation. *J Biol Chem* **281**: 35727–35734.

Lehle L, Strahl S, Tanner W (2006). Protein glycosylation, conserved from yeast to man: a model organism helps elucidate congenital human diseases. *Angew Chem Int Ed Engl* **45**: 6802–6818.

Lensen JF, Rops AL, Wijnhoven TJ, Hafmans T, Feitz WF, Oosterwijk E *et al*. (2005). Localization and functional characterization of gly-cosaminoglycan domains in the normal human kidney as revealed by phage display-derived single chain antibodies. *J Am Soc Nephrol* **16**: 1279–1288.

Lindahl U (2007). Heparan sulfate-protein interactions – a concept for drug design? *Thromb Haemost* **98**: 109–115.

Lindahl U, Backstrom G, Thunberg L, Leder IG (1989). Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin. *Proc Natl Acad Sci USA* **77**: 6551–6555.

Liu J, Shriver Z, Pope RM, Thorp SC, Duncan MB, Copeland RJ *et al*. (2002). Characterization of a heparin sulfate octasaccharide that binds to herpes simplex virus type 1 glycoprotein D. *J Biol Chem* **277**: 33456–33467.

Lortat-Jacob H, Turnbull JE, Grimaud JA (1995). Molecular organiza-tion of the interferon y-binding domain in heparan sulphate. *Biochem J* **310**: 497–505.

McKillop AM, McCluskey JT, Boyd AC, Mooney MH, Flatt PR, O'Harte FP (2000). Production and characterization of specific antibodies for evaluation of glycated insulin in plasma and biological tissues. *J Endocrinol* **167**: 153–163.

Merry CLM, Lyon M, Deakin JA, Hopwood JJ, Gallagher JT (1999). Highly sensitive sequencing of the sulfated domains of heparan sulfate. *J Biol Chem* **274**: 18455–18462.

Middleton J, Patterson AM, Gardner L, Schmutz C, Ashton BA (2002). Leukocyte extravasation: chemokine transport and presentation by the endothelium. *Blood* **100**: 3853–3860.

Mohammadi M, Olsen SK, Goetz R (2005). A protein canyon in the FGF–FGF receptor dimer selects from an a' la carte menu of heparan sulfate motifs. *Curr Opinion Struct Biol* **15**: 1–11.

Morimoto-Tomita M, Uchimura K, Werb Z, Hemmerich S, Rosen SD (2002). Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J Biol Chem* **277**: 49175–49185.

Mulloy B, Rider CC (2006). Cytokines and proteoglycans: an intro-ductory overview. *Biochem Soc Trans* **34**: 409–413.

Oscarsson LG, Pejler G, Lindahl U (1989). Location of the antithrombin-binding sequence in the heparin chain. *J Biol Chem* **264**: 296–304.

de Paz JL, Moseman EA, Noti C, Polito L, von Andrian UH, Seeberger PH (2007). Profiling heparin-chemokine interactions using synthetic tools. *ACS Chem Biol* **2**: 735–744.

Potzinger H, Geretti E, Brandner B, Wabitsch V, Piccinini AM, Rek A *et al*. (2006). Developing chemokine mutants with improved proteoglycan affinity and knocked-out GPCR activity as anti-inflammatory recombinant drugs. *Biochem Soc Trans* **34**: 435–437.

Proudfoot AE (2006). The biological relevance of chemokine-proteoglycan interactions. *Biochem Soc Trans* **34**: 422–426.

Proudfoot AEI, Souza ALS, Muzio V (2008). The use of chemokine antagonists in EAE models. *J Neuroimmunol* **198**: 27–30.

Prydz K, Dalen KT (2000). Synthesis and sorting of proteoglycans. *J Cell Sci* **113**: 193–205.

Robinson CJ, Mulloy B, Gallagher JT, Stringer SE (2006). VEGF165-binding sites within heparan sulfate encompass two highly sulfated domain and can be liberated by K5 lyase. *J Biol Chem* **281**: 1731–1740.

Roughley PJ (2006). The structure and function of cartilage proteogly-cans. *Eur Cell Mater* **12**: 92–101.

Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA (2001). Glyco-sylation and the immune system. *Science* **291**: 2370–2376.

Rydberg L (2001). ABO-incompatibility in solid organ transplantation. *Transfus Med* **11**: 325–342.

Saad OM, Myers RA, Castleton DL, Leary JA (2005). Analysis of hyalu-ronan content in chondroitin sulfate preparations by using selec-

tive enzymatic digestion and electrospray ionization mass spectrometry. *Anal Biochem* **344**: 232–239.

Sanderson RD, Yang Y, Kelly T, MacLeod V, Dai Y, Theus A (2005). Enzymatic remodeling of heparan sulfate proteoglycans within the tumor microenvironment: growth regulation and the prospect of new cancer therapies. *J Cell Biochem* **96**: 897–905.

Schenauer MR, Yu Y, Sweeney MD, Leary JA (2007). CCR2 chemokines bind selectively to acetylated heparan sulfate octasaccharides. *J Biol Chem* **282**: 25182–25188.

Shahrara S, Proudfoot AE, Park CC, Volin MV, Haines GK, Woods JM *et al.* (2008). Inhibition of monocyte chemoattractant protein-1 ameliorates rat adjuvant-induced arthritis. *J Immunol* **180**: 3447–3456.

Shriver Z, Raman R, Venkataraman G, Drummond K, Turnbull J, Toida T *et al.* (2000). Sequencing of 3-O sulfate containing heparin decasaccharides with a partial antithrombin III binding site. *Proc Natl Acad Sci USA* **97**: 10359–10364.

Shur BD, Rodeheffer C, Ensslin MA, Lyng R, Raymond A (2006). Identification of novel gamete receptors that mediate sperm adhesion to the egg coat. *Mol Cell Endocrinol* **250**: 137–148.

Spillmann D, Witt D, Lindahl U (1998). Defining the interleukin-8-binding domain of heparan sulfate. *J Biol Chem* **273**: 15487–15493.

Stringer SE, Gallagher JT (1997). Specific binding of the chemokine platelet factor 4 to heparan sulfate. *J Biol Chem* **272**: 20508–20514.

Stringer SE, Forster MJ, Mulloy B, Bishop CR, Graham GJ, Gallagher JT (2002). Characterization of the binding site on heparan sulfate for macrophage inflammatory protein 1α. *Blood* **100**: 1543–2703.

Taylor KR, Gallo RL (2006). Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *FASEB J* **20**: 9–22.

Tkachenko E, Rhodes JM, Simons M (2005). Syndecans: new kids on the signaling block. *Circ Res* **96**: 488–500.

Trowbridge JM, Gallo RL (2002). Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **12**: 117R–125R.

Turnbull JE, Hopwood JJ, Gallagher JT (1999). A strategy for rapid sequencing of heparan sulfate and heparin saccharides. *Proc Natl Acad Sci USA* **96**: 2698–2703.

Turnbull JE, Fernig DG, Ke Y, Wilkinson MC, Gallagher JT (1992). Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate. *J Biol Chem* **267**: 10337–10341.

Van Acker GJ, Perides G, Steer ML (2006). Co-localization hypothesis: a mechanism for the intrapancreatic activation of digestive enzymes during the early phases of acute pancreatitis. *World J Gastroenterol* **12**: 1985–1990.

Vanpouille C, Deligny A, Delehedde M, Denys A, Melchior A, Liénard X *et al.* (2007). The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue. *J Biol Chem* **282**: 24416–24429.

Varki A (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**: 97–130.

Vlodavsky I, Goldshmidt O, Zcharia E, Atzmon R, Rangini-Guatta Z, Elkin M *et al.* (2002). Mammalian heparanase: involvement in cancer metastasis, angiogenesis and normal development. *Semin Cancer Biol* **12**: 121–129.

Vlodavsky I, Ilan N, Naggi A, Casu B (2007). Heparanase: structure, biological functions, and inhibition by heparin-derived mimetics of heparan sulfate. *Curr Pharm Des* **13**: 2057–2073.

Weber C, Koenen RR (2006). Fine-tuning leukocyte responses: towards a chemokine 'interactome'. *Trends Immunol* **27**: 268–273.

Whelan SA, Hart GW (2003). Proteomic approaches to analyze the dynamic relationships between nucleocytoplasmic protein glycosylation and phosphorylation. *Circ Res* **93**: 1047–1058.

Yang XD, Corvalan JRF, Wang P, Roy CMN, Davis CG (1999). Fully human anti-interleukin-8 monoclonal antibodies: potential therapeutics for the treatment of inflammatory disease states. *J Leukoc Biol* **66**: 401–410.

Yu Y, Sweeney MD, Saad OM, Crown SE, Hsu AR, Handel TM *et al.* (2005). Chemokine-glycosaminoglycan binding: specificity for CCR2 ligand binding to highly sulfated oligosaccharides using FTICR mass spectrometry. *J Biol Chem* **280**: 32200–32208.

Yu Y, Sweeney MD, Saad OM, Leary JA (2006). Potential inhibitors of chemokine function: analysis of noncovalent complexes of CC chemokine and small polyanionic molecules by ESI FT-ICR mass spectrometry. *J Am Soc Mass Spectrom* **17**: 524–535.

Zamfir A, Seidler DG, Schonherr E, Kresse H, Peter-Katalinic J (2004). On-line sheathless capillary electrophoresis/nanoelectrospray ionization-tandem mass spectrometry for the analysis of glycosaminoglycan oligosaccharides. *Electrophoresis* **25**: 2010–2016.