Remodeling of heparan sulfate : functional and structural characterization of human endosulfatase HSulf-2

Rana El Masri

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de
DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES
Spécialité: **Nanobiologie**
Arrêté ministériel : 25 mai 2016

Présentée par
« Rana El Masri »

Thèse dirigée par **« Romain Vivès », DR2, CNRS**

préparée au sein du groupe **Structure et Activité des Glycosaminoglycanes à l'Institut de Biologie Structurale**
dans l'**École Doctorale Chimie et Sciences du Vivant**

Remodeling of Heparan Sulfate: Functional and structural characterization of human endosulfatase HSulf-2

Thèse soutenue publiquement le **20 septembre 2019**,
Devant le jury composé de :

Dr. Pascale MARCHOT
Directeur de recherche, AFMB Marseille, Rapporteur

Dr. Kenji UCHIMURA
Directeur de recherche, UGSF Lille, Rapporteur

Dr. Régis DANIEL
Directeur de recherche, LAMBE Evry, Examinateur

Dr. Nicole THIELENS
Directeur de recherche, IBS Grenoble, Président

Dr. Romain VIVES
Directeur de recherche, IBS Grenoble, Directeur de thèse

Table of Contents

Table of Figures

Table of Abbreviations and Acronyms

Anti-thrombin III: ATIII

BMP: Bone Morphogenic Protein

CS: Chondroitin Sulfate

DS: Dermatan Sulfate

ECM: Extracellular Matrix

EGF(R): Epidermal Growth Factor (Receptor)

ER: Endoplasmic Reticulum

EXT: Extostosin

FGF(R): Fibroblast Growth Factor (Receptor)

FGly: Formylglycin

GAG: Glycosaminoglycan

Gal: Galactose

GalNAc: N-acetyl galactosamine

GDNF: Glial cell-derived neurotrophic factor

GlcA: Glucoronic Acid

GlcNAc: N-acetyl glucosamine

HA: Hyaluronan

HAS: Hyaluronan Synthase

HEK: Human Embryonic Kidney

HGF: Hepatocyte Growth Factor

HH: Hedgehog

HIV: Human Immunodeficiency Viruses

HS: Heparan Sulfate

HSV: Herpes Simplex Virus

HP: Heparin

IdoA: Iduronic Acid

IFN: Interferon

IL: Interleukin

KS: Keratan Sulfate

MAPK: Mitogen-activated protein kinases

MMP: Matrix Metalloproteinase

NDST: *N*-deacetyl sulfotransferase

NMR: Nuclear Magnetic Resonance

OST: *O*-Sulfotransferase

PAPS: 3'-Phosphoadenosine 5'-Phosphosulfate

PDGF(R): Platelet Derived Growth Factor (Receptor)

PG: Proteoglycan

PNN: Perineuronal Net

RT-PCR: Reverse transcription polymerase chain reaction

SHH: Sonic Hedgehog

SLRP: Small Leucin Rich Proteoglycan

SPR: Surface Plasmon Resonance

TGFβ: Transforming growth factor beta

UDP: Uridine Diphosphate

VEGF(R): Vascular Endothelial Growth Factor (Receptor)

Wnt: Wingless type

Xyl: Xylose

4MUS: 4-methylumbelliferyl sulfate

CHAPTER

I

Introduction

This first chapter lays the background on the topics that have been investigated during the course of this PhD project.

As this manuscript focuses on Sulf endosulfatase enzymes, it is important to provide an overview about the environment, in which these enzymes act. Sulfs are unique extracellular endosulfatases that control the structure of heparan sulfate (HS), a type of glycosaminoglycan (GAG) polysaccharide, located within the extracellular matrix (ECM) and at the cell surface, borne by specific types of proteoglycans (PGs). Modifying the structure of HS affects its ability to interact with its ligands, resulting in critical functional consequences. Sulfs are thus implicated in important physiological and pathological processes. In this introduction, I will first introduce the PGs and the GAGs: I will discuss their different types, diversity, roles and their highly complex biosynthesis process. I will then focus on the ability of HS to interact with a broad panel of ligands and its role in the regulation of major cellular functions. Finally, I will present the sulfatase family in general to focus then on Sulfs and on the features that make them unique amongst sulfatases. I will summarize the state of the art about the Sulfs, regarding their mechanisms of action, their critical control of HS activity and their implications in physiopathological processes especially in cancer.

1. Proteoglycans and Glycosaminoglycans

The life of pluricellular organisms is largely dependent on cells communicating with each other, and with their surroundings. This interaction is mediated by a broad range of soluble signaling proteins such as growth factors, interleukins, cytokines, chemokines…that elicit signaling pathways by binding to their cognate receptors in order to trigger specific cell responses. In physiological conditions, these diffusible proteins are released at very low concentrations and their activities are highly regulated in time and space. To access to their receptors, most of these proteins must come across the glycocalyx, a thick layer of glycosylated molecules found at the cell surface, and in the ECM, a solid substrate of macromolecules that ensures tissue cohesiveness. Within these two extracellular compartments, PGs play a central role in controlling the diffusion and the activity of most of these signaling proteins.

PGs are therefore implicated in most major cellular processes, such as cell proliferation, differentiation, migration, adhesion, chemoattraction, inflammation, immune responses, control of angiogenesis and coagulation… (Iozzo and Schaefer, 2015; Sarrazin et al., 2011). In addition, PGs provide cell surface attachment sites for a variety of pathogen micro-organisms, such as parasites (Plasmodium falciparum, Toxoplasma gondii…), bacteria (Streptococcus pneumoniae, Helicobacter…) and viruses (HIV, HSV, Hepatitis C…) (Bartlett and Park, 2010). These activities depend on the glycan component of the PGs.

Structurally, PGs are composed of a protein core, to which linear, anionic polysaccharides that belong to the GAG family are covalently attached. One PG can be decorated with many GAG chains, possibly of different types.

GAGs comprise hyaluronan (HA), chondroitin sulfate (CS), dermatan sulfate (DS), keratan sulfate (KS), heparin (HP) and heparan sulfate (HS). HA is the only GAG that is not associated to a PG protein core and is thus found as free chains.

1.1. Structure and classification of PGs

Several scientific studies led to the discovery of PGs. The first one was the demonstration of a covalent linkage between a serine and a CS chain (Muir, 1958). Later, protein/polysaccharide complexes were isolated from cartilage by ultracentrifugation, but their exact nature could not be determined (Pal et al., 1966). Finally, a tetrasaccharide linker

GlcA-Gal-Gal-Xyl that connects the serine amino-acid residue to GAG polysaccharide was identified (Rodén and Smith, 1966).

Both protein and sugar moieties have important contribution in the biological properties of PGs. In fact, the lack of a particular PG protein core or of GAG biosynthesis enzymes in eukaryotic organisms leads to diverse severe phenotypes (Lindahl, 2014). Indeed, protein cores are known to determine the localization of PG and also to transduce signals through cell membranes (Alexopoulou et al., 2007), while GAG chains elicit their biological functions through their ability to interact with a variety of signaling proteins, controlling thus their bioavailability, access to their receptors, storage, and protection from proteolysis… In addition, it has been shown for some PGs such as HSPG perlecan, that biological functions can be triggered by the ability of their protein core to interact with other protein ligands, including growth factors or surface receptors in order to stabilize the cell – matrix cohesion (Whitelock et al., 2008).

Nowadays, around 50 distinct mammalian genes encoding for PG protein cores have been discovered, including variants resulting from alternative splicing. PGs can be classified depending on their localization: they can be extracellular, membrane bound, or intracellular (Figure 1). Until now, only one PG appears to be **intracellular**, the serglycin (Kolset and Pejler, 2011). Regarding **membrane bound** PGs, they generally feature HS chains. Their major functions are the control of the access of growth factors to their receptors and the regulation of morphogens to maintain their gradients during embryogenesis and regenerative processes (Iozzo and Schaefer, 2015). The two major families of HSPGs are Syndecans and Glypicans. HSPGs are either transmembrane proteins associated to the membrane *via* their protein core (Syndecans) or *via* a glycosyl-phosphatidyl-inositol (GPI) anchor (Glypicans). Other membrane bound PGs include Betaglycan, CSPG4/NG2 and Phosphacan, the latter two being CSPGs. HSPGs can be also attached to the **pericellular matrix** *via* integrin or other receptors, playing important roles in maintaining the basement membrane by interacting with each other and with the other constituents such as laminins, collagen type IV, and nidogen (Iozzo and Schaefer, 2015). This group is constituted of 4 PGs: Perlecan and its shorter form: the Endorepellin, Agrin, Collagens XVIII and XV. Collagen XV lacks the HS chain, which is substituted with a CS chain.

PGs located in the **extracellular matrix** are the largest family and are mostly CSPGs and DSPGs. They are divided in three groups: the first one includes Hyalectans (hyaluronan and lectin binding PGs) and is composed of 4 members: Agrecan, Versican, Neurocan and

Brevican (Iozzo and Murdoch, 1996). They bind to hyaluronan and form enormous complexes of high viscosity, playing structural roles in cartilage, blood vessels and central nervous system. The second one gathers the Small leucine-rich proteoglycans (SLRPs) comprising 18 members (Neill et al., 2015). As their name indicates, they are constituted of small proteins composed of leucine rich repeat regions (LRR). They are ubiquitously expressed in most ECM and play important roles in the embryonic development and homeostasis. They are implicated in processes such as migration, proliferation, autophagy, apoptosis, innate immunity and angiogenesis, due to their interaction with many receptor tyrosine kinases (RTKs) and Toll-like receptors. They can be found in circulation as well. SLRPs are divided in 5 classes. The majority of class I harbours CS and DC chains, class II features KS chains and class III one of both or no GAGs. Class IV and V lack GAG chains except for one. Finally, the third group corresponds to the SPOCK family. They are composed of 3 HSPGs that bind to calcium and are called Testicans (Iozzo and Schaefer, 2015).

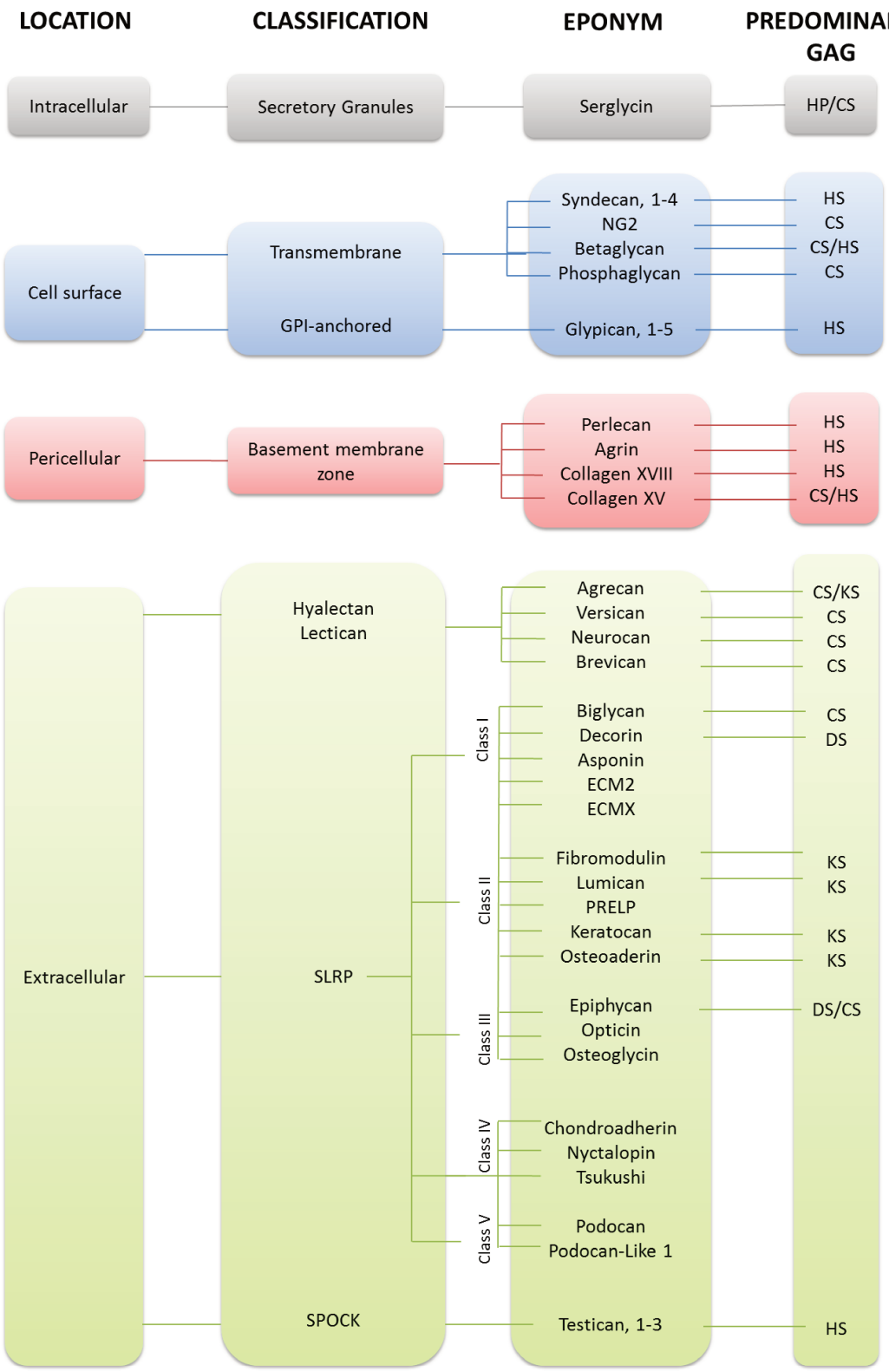


Figure 1: Classification of PGs depending on their localization. Adapted from (Iozzo and Schaefer, 2015).

1.1.1. Intracellular HSPGs: Serglycin

Serglycin is the only PG that appears to be intracellular. It was the first PG identified at cDNA level and was given this name because of the high occurrence of Serine - Glycine

repeats, the GAG attachment sites. Another property making this PG unique is its ability to adopt different structures, depending on the cell type. It can display CS chains, HS chains, and it is also the only PG that can feature HP chain (Kolset and Pejler, 2011). Serglycin can also exist as an hybrid PG containing different types of GAG chains (Lidholt et al., 1995).

When expressed in connective tissues like the granules of mast cells, Serglycin is rich in sulfated HP. Serglycin is also present in the intracytoplasmic granules of other inflammatory cells, such as lymphocytes, monocytes… (Kolset and Gallagher, 1990). In this case, it is associated to CS chains. Interestingly, when these cells are found in circulation, the CS chains are low sulfated, but once the cells are activated, they tend to have higher sulfation levels. Notably, Serglycin was also shown to contain HS chains in primary murine macrophages (Iozzo and Schaefer, 2015). In addition, Serglycin can be found in non-immune cells, including endothelial cells, smooth muscle cells and chondrocytes (Lemire et al., 2007; Meen et al., 2011; Zhang et al., 2010). It has been shown that the synthesis and secretion of Serglycin are increased following activation of endothelial cells. Finally, Serglycin expression is induced in myeloma cells and its overexpression enhances their metastatic potential (Theocharis et al., 2006).

The close proximity of Ser-Gly repeats within the protein core leads to formation of sulfated HP clusters. This enables the Serglycin to trap various compounds such as proteases, chemokines, histamine, serotonin in the granules of cells *via* electrostatic interactions between the negative charges of HP and the basic residues of these compounds (Kolset and Pejler, 2011). This storage is impaired by the knockout (KO) of Serglycin or by the KO of NDST-2 (a biosynthesis enzyme responsible of HS/HP *N*-sulfation, see page 33) in a similar way (Forsberg et al., 1999). The various stored components can be released in complexes with Serglycin to achieve specific functions. For example, Serglycin regulates cell apoptosis by releasing Serglycin-protease complexes from the granules to the cytosol of damaged cells (Braga et al., 2007). Furthermore, Serglycin is crucial to optimize the presentation of proteases to their substrates in the ECM (Humphries et al., 1999). It can also facilitate the transport of chemokines to their target cells and modulate their activities (Meen et al., 2011; Wagner et al., 1998). It is worth noting that after secretion of the complexes, histamine and chemokines can detach from Serglycin (Kolset and Pejler, 2011) (Figure 2).

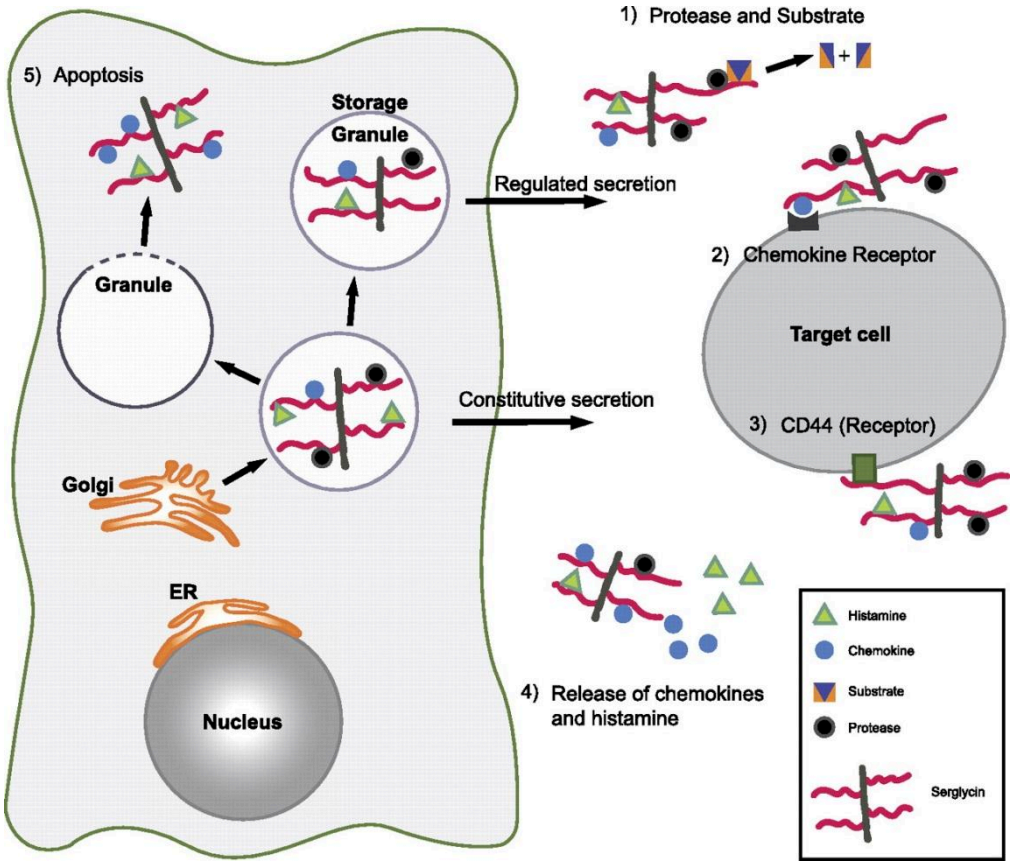


Figure 2: Functions of serglycin. From (Kolset and Pejler, 2011).

1.1.2. Membrane HSPG: Syndecans and Glypicans

Syndecan family is composed of 4 proteins (Syndecan 1-4). They all contain an ectodomain, a transmembrane domain and an intracellular domain. The latter two domains are well conserved amongst Syndecans. In contrast, only 10-20% of the ectodomain amino-acid sequence is conserved (Iozzo and Schaefer, 2015). In Syndecans, up to 5 GAGs chains can be found attached to the ectodomain of the protein, which can be either HS or CS. This domain is natively disordered, which allows the interaction of Syndecans with various ligands, and their implications in major biological functions (Leonova and Galzitskaia, 2013). The transmembrane domain is characterized by its ability to elicit homo or hetero dimerization (Teng et al., 2012). The intracellular domain is composed of 2 conserved regions separated by a variable amino-acid sequence, and of a conserved peptide signature that binds to PDZ containing proteins (Figure 3). PDZ containing proteins are known for their role in anchoring transmembrane receptor proteins to the cytoskeleton, organizing thus large signaling complexes. Syndecans are characterized by their involvement in the uptake of exosomes and by acting as endocytosis receptors for the clearance of bound ligands like lipoproteins (Stanford et al., 2009). This mechanism is important for delivering nutrients to

cells or for the removal of bioactive factors from the environments. They can also be co-receptors for many TKRs like growth factor receptor kinases playing role in the modulation of the duration of signaling reactions (Carey, 1997).

At the cell surface, Syndecans can undergo protein "shedding" (see page 47). This process consists of the cleavage of the extracellular domain of Syndecan, releasing thus soluble HS-containing protein fragments that have important biological functions (Teng et al., 2012).

Glypican family is composed of 6 members with similar protein cores. Interestingly and in contrast with Syndecans, only HS chains are attached near the juxtamembrane region, covering thus a large part of the plasma membrane. This may facilitate the presentation of signaling proteins to their receptors, such as Hedgehog (HH), Wingless type (Wnt), and Fibroblast Growth Factors (FGFs) (Capurro et al., 2005, 2008). Glypicans can feature 1 to 3 HS chains. Glypicans can also be shedded by proteases like furin-like convertases, leading to the release of the ectodomain, or also by lipases, which releases the entire protein core (Figure 3). This process may modulate the regulation of morphogen gradients by Glypicans (Filmus and Capurro, 2014).

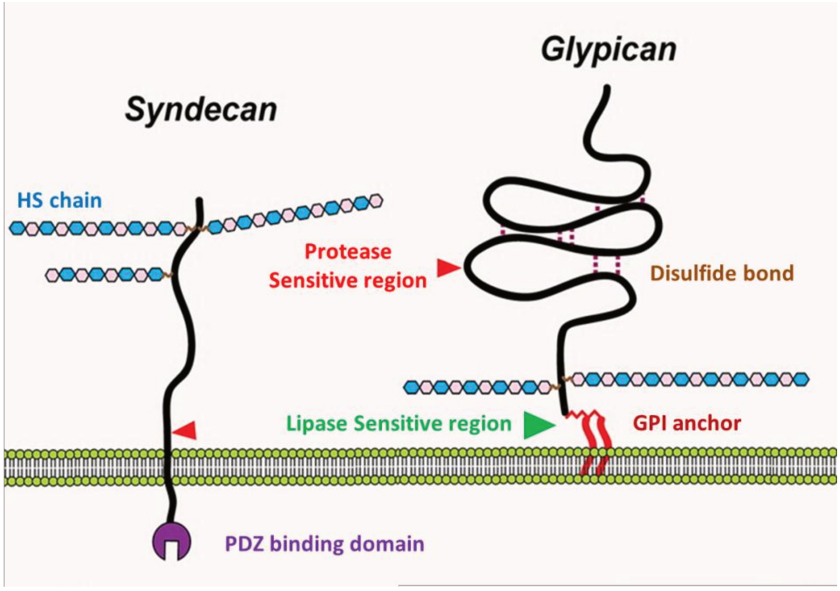


Figure 3: Schematic representation of HSPG Syndecan and Glypican. Adapted from (Iozzo and Schaefer, 2015).

1.1.3. Extracellular CSPG: Decorin

Decorin is a CSPG that belongs to class I of SLRPs. Its name was given because of its ability to bind and to decorate collagen fibrils. It has a banana shape with an N-terminal

region featuring one CS or DS chain and a concave central region containing 12 LRRs. The C-terminal part of the protein seems to have a critical role, given that a truncated Decorin, lacking the C-terminal 33 residues, called the "ear repeat", was identified in patients with congenital stromal corneal dystrophy (Figure 4) (Bredrup et al., 2005). Decorin is known to maintain structural integrity of many organs and to regulate physiological and pathological processes, through its interaction with many signaling proteins (Gubbiotti et al., 2016). The binding of Decorin to its ligands can involve either the GAG chain, or the protein core. Indeed, Decorin binds to collagen I and RTKs like vascular endothelial growth factor receptor 2 (VEGFR2) and epidermal growth factor receptor (EGFR) *via* specific LRR regions, thus resulting in activation of downstream signaling cascades. Indeed, Decorin can induce autophagy in endothelial, epithelial and glioma cells. It can also activate mitophagy in cancer cells (Buraschi et al., 2019). Interestingly, Decorin has the capacity to dimerize, thereby hiding its core protein (Islam et al., 2013). This can prevent its activity, and thus perturb many signaling pathways. It was shown that the dimerization of decorin is reversible, given to Decorin the ability to altern between both forms depending on the environment needs (Islam et al., 2013).

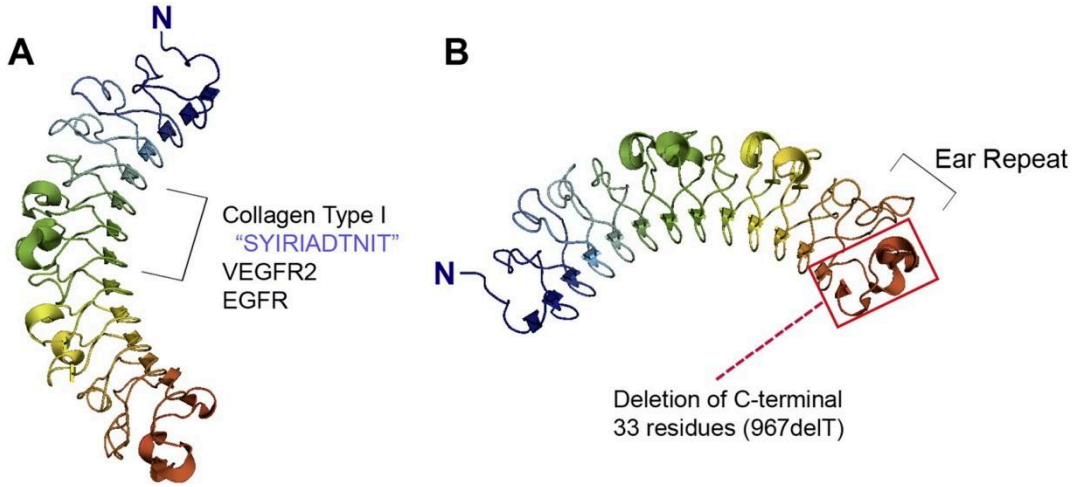


Figure 4: Graphical representation of the structure of Decorin. From (Gubbiotti et al., 2016).

1.2. Structure and classification of GAGs

GAGs are linear and highly heterogeneous complex polysaccharides, constituted by a repetition of disaccharide units (n), composed of an alternating hexosamine and uronic acid or galactose (Gal) that differ according to the GAG type. The geometry of linkage between the sugar units may also vary (α or β). The uronic acid can be either glucuronic acid (GlcA) or its C5 epimerized form, the iduronic acid (IdoA). Regarding the hexosamine, it can be

either *N*-acetyl glucosamine (GlcNAc) or *N*-acetyl galactosamine (GalNAc). The structural diversity of GAG chains is enhanced by the possible modifications of these units. Indeed, the amine group in the hexosamine can be *N*-acetylated or *N*-sulfated and the hydroxyl groups in the position C2 of the uronic acids and C3, C4 and C6 position of hexosamine can be *O*-sulfated as well (Figure 5). All these types of modification confer to GAGs their strong anionic properties and huge structural diversity. For example, a simple octasaccharide (n=4) can represent over millions different sequences (Gandhi and Mancera, 2008).

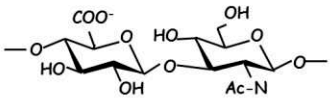
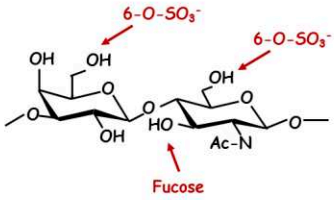
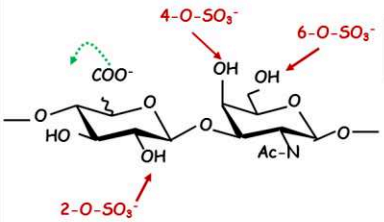
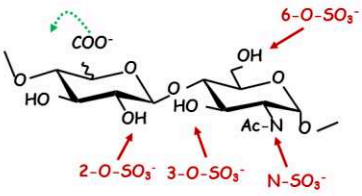
Glycosaminoglycan type	Disaccharide units and the possible modifications	Disaccharide repetition number (n) and Molecular weight (MW)
Hyaluronan	GlcA-β(1→3) – GlcNAc-β(1→4) 	n = 8 – 16000 MW = 4 – 8000 KDa
Keratan Sulfate	Gal-β(1→4) – GalNAc-β(1→3) 	n = 9 – 45 MW = 4 – 20 KDa
Chondroitin Sulfate Dermatan Sulfate	GlcA-β(1→3) - GalNAc-β(1→4) IdoA-α(1→3) - GalNAc-β(1→4) 	n = 10 – 130 MW = 5 – 60 KDa
Heparan Sulfate/Heparin	GlcA-β(1→4) – GlcNAc-α(1→4) IdoA-α(1→4) – GlcNAc-α(1→4) 	n = 20 – 100 MW = 10 – 50 KDa

Figure 5: Structure and composition of different types of GAG. Adapted from (Gandhi and Mancera, 2008; Pomin and Mulloy, 2018).

1.2.1. Hyaluronan

The disaccharide of hyaluronan (HA) is a GlcA β(1→3) linked to a GlcNAc. HA is remarkable as being the only GAG that is not associated to a protein core and does not undergo sulfation modifications. Furthermore, unlike other GAGs that are assembled in the Golgi lumen, HA polymerization occurs at the inner face of the plasma membrane by 3 different HA synthases (HAS) that recruit uridine diphosphate (UDP) sugars from the cytoplasm. HAS differ by their expression in time and space and by their ability to produce HA with different lengths (Weigel et al., 1997). HA is found abundantly in the eye vitreous humor and in connective tissues. It has also been suggested that HA may be found inside the cell but its function there is not well understood (Almond, 2007).

As mentioned before, it has been shown in some tissues like cartilage and brain that HA can act as binding partners for PGs called Hyalectans. HA can also bind to cell receptors such as the lymphocyte homing receptor CD44. These interactions contribute to the formation of large aggregated networks that determine the physical form of the tissue, and that play a major role in tissue hydration by capturing water molecules. In addition, they imply HA in important biological processes like regeneration, embryogenesis, cell motility, inflammation and angiogenesis (Almond, 2007).

Thanks to these biological properties and being non-toxic and non-immunogenic, HA is widely used in the development of engineered tissues and biomaterials in the biomedical field, with esthetic, orthopedic, cardiovascular, pharmacologic and oncologic applications (Allison and Grande-Allen, 2006). For example, it can be used for drug delivery in eye surgery or in wound repair.

The turnover of HA is very rapid (Laurent and Reed, 1991). It can occur by lymphatic removal and degradation in lymph nodes and liver. It is suggested that there are scavengers receptors on the liver endothelial cell surface implicated in that process (McCourt, 1999). It can also occur by hyaluronidase enzymes that degrade HA into smaller fragments. Several degrading enzymes have been identified in bacteria and fungi. These HA fragments seem to play important biological functions. They cause changes in tissue morphology, stimulate fibroblast proliferation and collagen synthesis, and induce cytokine production in dendritic cells (Stern and Jedrzejas, 2006).

1.2.2. Keratan sulfate

Unlike other GAGs, keratan sulfate (KS) does not contain uronic acid. It is mostly composed of a Gal-GlcNAc repeat. Both saccharides, mostly the GlcNAc, can be 6-*O*-sulfated. The biosynthesis of KS initiates differently among KS types. Indeed, KS can be attached to PG protein core *via N*- or *O*-linkages. For some KS, called KS I, the initiation occurs with the formation of *N*-linkage between GlcNAc and an asparagine of the protein core. For other KS, called KS II, it involves *O*-linkage between a GalNAc and a serine/threonine. Finally, a third group is initiated by an *O*-linkage between a mannose and a serine (Figure 6). Following initiation, KS chain elongation involves 2 transferases: the β-1,3-*N*-acetylglucosaminyl transferase (β3GnT) and β-1,4-galactosyl transferase (β4GalT-1). The polysaccharide maturation depends on 2 sulfotransferases: *N*-acetylglucosaminyl-6-sulfotransferase (GlcNAc6ST) and KS galactosyl sulfotransferase (KSGalST). This process results in the formation of a wide range of KS, with diverse protein linkages, chain lengths and sulfation degrees (Caterson and Melrose, 2018). The sulfation is variable and critical in some tissues. Indeed, the alteration of KS sulfation degree in human leads to corneal opacity in macular corneal dystrophy (Dang et al., 2009). Moreover, highly sulfated KS has been associated with a number of tumors. In addition, a recent study showed that the contribution of GlcNAc6ST isoforms in brain is stage dependent and that in adult brains, GlcNAc6ST3 plays a very important role in the synthesis of perineuronal net (PNN) components (Narentuya et al., 2019) (PNNs are highly conserved ECM structure found around neurons in the central nervous system).

Another specific structural feature of KS is the presence of L-fucose that modifies the GlcNAc and *N*-acetylneuraminic acids at the non-reducing end of their chains, called capping structure (Figure 6). Interestingly, the presence of these structures prevents the degradation of KS by keratanase I, keratanase II and endo-β-D-galactosidase (Caterson and Melrose, 2018).

KS is widely distributed over the human body. The richest tissue is the cornea, where KS I is mostly found, then in the brain where KS III and highly sulfated KS are present. It can also be found in other connective tissues such as cartilage and bone, where KS II is abundant, in epithelial tissues and in central and peripheral nervous system (Funderburgh, 2002).

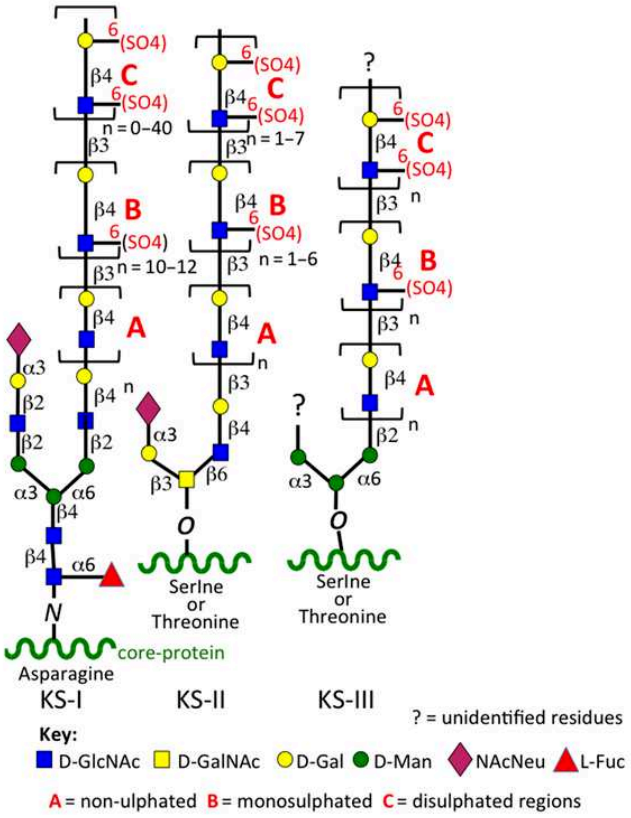


Figure 6: Schematic representation of different types of KS. From (Caterson and Melrose, 2018).

Unlike CS and HS, only few studies have reported the interaction of KS with signaling proteins. It has been shown that corneal KS interacts with sonic hedgehog (SHH) and FGFs, and also to a number of kinases, cytoskeletal components and soluble proteins implicated in tissue development (Conrad et al., 2010; Weyers et al., 2013). In addition, KS plays critical roles in neurotransmission and nerve regeneration, being able to bind to various nerve regulatory proteins, like semaphorin, nerve growth factors, slit, robo and ephrin (Conrad et al., 2010). In neural tissues, some of these interactions induce actin depolymerization, cytoskeletal re-organization and cell signaling.

1.2.3. Chondroitin Sulfate/Dermatan sulfate

The disaccharide repeating unit of CS is composed of a GalNAc linked to a GlcA *via* β(1→4) linkage. Both sugars are subject to sulfation at different positions, giving rise to various types of CS disaccharides, which presence confers selective functions to the overall polysaccharide. The most common sulfation patterns are represented as follow. Regarding the GalNAc, the addition of a sulfate group at position C4 leads to CS-A disaccharide type, while its addition at position C6 corresponds to CS-C. Sulfation at both positions yields to 4-*O*/6-*O*-

sulfated CS-E disaccharides. A 6-*O*-sulfated GalNAc and a 2-*O*-sulfated GlcA, results in CS-D disaccharides. Finally, the last disaccharide type called CS-B or dermatan sulfate (DS), occurs from the epimerization of CS-A unit (GlcA into L-iduronic acid), becoming the most diverse type (Figure 7, Djerbal et al., 2017). Noteworthy, one CS chain contains different disaccharides unit types, but is referred by the most abundant one.

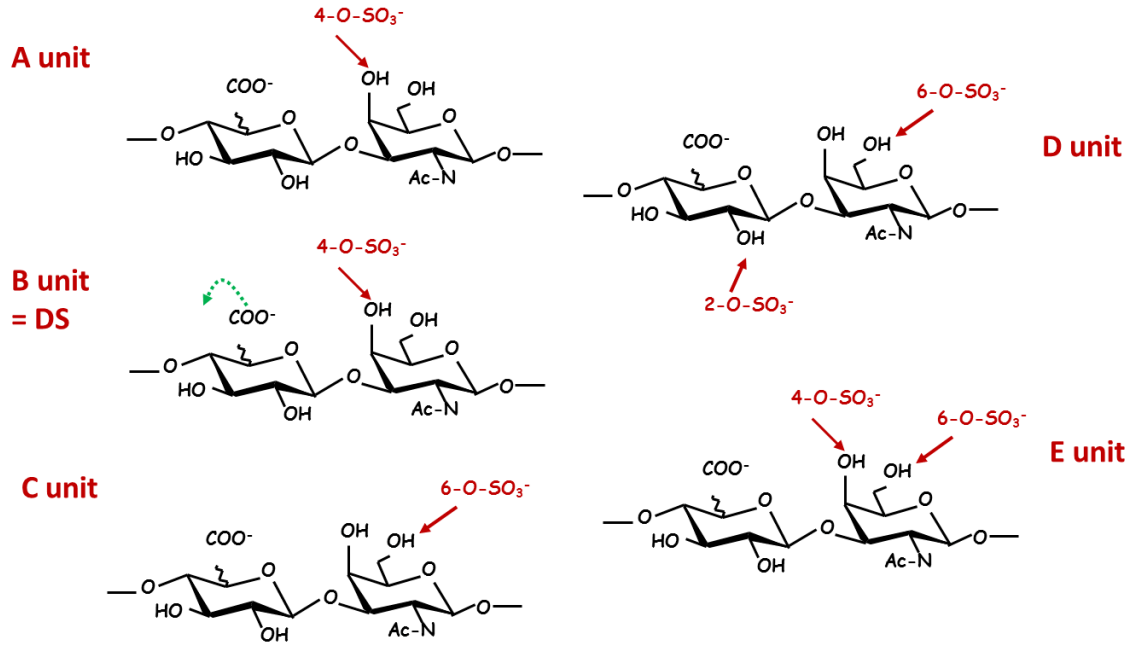


Figure 7: Structure of most common disaccharide units of chondroitin sulfate. Adapted from (Djerbal et al., 2017).

CS is ubiquitously present in the ECM of tissues. It is mostly abundant in cartilage, where it controls the resistance and elasticity of the tissue and it regulates important processes. For example, it has been shown that CS inhibits various inflammatory enzymes and cytokines. It also downregulates chondrocyte apoptosis and inhibits metalloproteinase degradation (Henrotin et al., 2010). CS is also highly present in the central nervous system, where it controls the development and the maintenance of brain plasticity and function. It is implicated in the neuronal cell migration, the axon regeneration and the memory retention control. These functions take place through the interactions between CS and various proteins, including guidance proteins such as semaphorins, growth factors, receptors or adhesion molecules (Laabs et al., 2005).

Like all GAGs, there is a tight relationship between CS functions and structures, which differs depending on the development stage and the pathological conditions. Indeed, the

composition of CS is spatio-temporally controlled. For instance, the most expressed CS in nervous system is CS-C during embryogenesis, while it is CS-A in adulthood (Kitagawa et al., 1997). In addition, CS-A is the richest CS in the adult brain. However, CS-D, CS-E and CS-B, are more present in the PNNs (Deepa et al., 2006). Hence, CS-D contributes in the maintenance of naïve T lymphocytes in the spleen of young mice (Uchimura et al., 2002). Furthermore, only DS is able to interact with heparin cofactor II and with hepatocyte growth factor (HGF) (Lyon et al., 1998; Mascellani et al., 1993). Moreover, semaphorin 3A binds to CS-E and DS, but not to CS-D, given that CS-E and CS-D both have 2 sulfate groups (same charge over mass ratio) differing in the position (Dick et al., 2013). Finally, it has been shown that CS structure is modified and that the CS ratio in 6S/4S is reduced in cartilage of osteoarthritis patients, suggesting CS supply as a therapy for patients with cartilage damage (Henrotin et al., 2010).

1.2.4. Heparan sulfate/Heparin

HS is ubiquitously expressed at cell surface and in ECM, while HP is expressed in mast cells. HP was first discovered for its anticoagulant effect and is now used in pharmacology as an anticoagulant drug in the treatment of thromboembolic diseases. HS and HP have a closely related structure. They are composed of a repetition of a GlcA β(1→4) linked to a GlcNAc. The disaccharides can be subject to series of modifications: the amino sugar can be deacetylated and subsequently *N*-sulfated, the GlcA can be epimerized into IdoA, and saccharide units can be variably substituted with *O*-sulfate groups, at the C6 (and occasionally C3) of glucosamine and at C2 of IdoA acid residues. These modifications occur in specific regions of the polysaccharide termed NS domains. HS and HP are thus both organized in an alternation of hypervariable and highly sulfated NS domains, and homogeneous, non or low sulfated regions termed the NAc domains, separated by intermediate NAc/NS transition zones (Figure 8). Nevertheless, HP (an average of 2.3 sulfate per disaccharide) are more sulfated compared to HS (an average of 0.8 sulfate per disaccharide) and features more extended NS domains with higher proportion of IdoA and 2-*O*-sulfation, (and higher 3-*O*-sulfation), in contrast to HS that is richer in GlcA (Sarrazin et al., 2011).

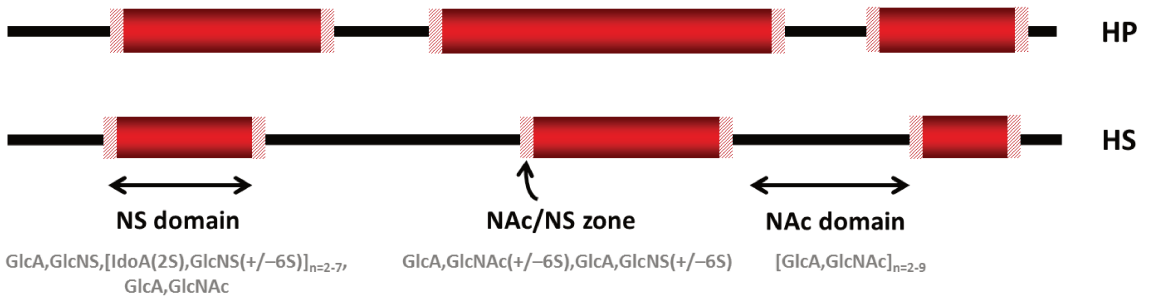


Figure 8: Structure of HP and HS. HP and HS consist of repeating GlcA-GlcNAc disaccharide units. In regions termed NS domains (red boxes), this motif is extensively modified by a series of modifications including epimerization and sulfation reactions. The overall structure of the polysaccharide chain thus exhibits an alternation of NS domains and non-modified regions (NAc domains, black line), separated by intermediate NAc/NS transition zones (dashed boxes).

This structural diversity results, theoretically, in the creation of 48 different disaccharides, making HP and HS the most complex GAGs, capable to interact with a wide variety of ligands. HS biological properties are thus largely dictated by their structure and the sulfation patterns of NS domains, which comprise the binding sites for most HS ligands (Capila and Linhardt, 2002; Imberty et al., 2007; Lindahl and Li, 2009). The degree of structural specificity of these binding sites varies from one ligand to another, but can be very high for some proteins, depending on the nature of the sulfates involved. For instance, rare 3-*O*-sulfates have been associated with highly selective interactions like with anti-thrombin III (ATIII), while abundant *N*-, 2-*O*-sulfates are part of most protein binding sites (Li and Kusche-Gullberg, 2016; Lindahl and Li, 2009). Finally, 6-*O*-sulfation may occupy an intermediate status, with clear indications of distinctive features (El Masri et al., 2017). As the first modification reaction, *N*-sulfation has been shown to be critical for determining NS domain size and distribution, and 2-*O*-sulfation influences the GlcA/IdoA epimerization ratio. These early modification steps may thus provide common structural and functional properties to the polysaccharide. In agreement with this, recent studies have demonstrated that *N*-deacetylase sulfotransferases (NDSTs), C5-epimerase and 2-*O*-sulfotransferase (2OST) (biosynthesis enzymes of HS) respectively responsible for HS *N*-sulfation, C5-epimerization and 2-*O*-sulfation, show processive activities (Carlsson et al., 2008; Préchoux et al., 2015). And earlier work on the biochemical characterization of HS reported that NS domains were composed of relatively homogeneous blocks of contiguous [GlcNS, IdoA(2S)] disaccharides (Merry et al., 1999; Pye et al., 1998). In contrast, the addition of 6-*O*-(and 3-*O*-) sulfates is much more heterogeneous. These late modifications may thus provide further diversity, and therefore potential specificity/selectivity, to the polysaccharide. In addition, it is worth noting

that 6-*O*-sulfation is the only type of sulfation to be significantly found outside NS domains, with almost half of the 6-*O*-sulfates present in the NAc/NS transition regions. It is also the only modification step to be regulated through both biosynthesis by 6-*O*-sulfotransferases (6OST) (see page 41) and post-synthesis processes by Sulfs that target specifically the NS domains (see page 62).

2. Biosynthesis of HP/HS and CS/DS

The biosynthesis of GAG chains is a complex process involving many enzymatic systems that takes place in the Golgi apparatus. It is composed of 3 main steps: the chain initiation, elongation and maturation. The initiation is a common step to the synthesis of HS and CS chain. It consists in the assembly of a tetrasaccharide linker on the protein core, which has been previously synthesized in the endoplasmic reticulum (ER) and transferred to the Golgi apparatus. Subsequent elongation and maturation steps both involve many enzymes that differ between HS and CS, such as glycosyltransferases, sulfotransferases and epimerases.

2.1. Initiation

The synthesis is initiated by the transfer of the xylose to specific serine residues of the dipeptide Ser-Gly of the protein core, the glycine being not essential for the GAG synthesis (Huber et al., 1988). The transfer is catalyzed by two xylosyltransferases (XylT-I and XylT-II) using UDP-xylose as a donor, and takes place in the ER or in the cis Golgi compartment, depending on the cell type.

The xylosylation process can be incomplete in some PG, meaning that not all the potential attachment sites are likely to bear GAG chains (Silbert and Sugumaran, 1995). This reaction is dependent on the amino-acid residues located around the dipeptide, and also on the serine itself. Post-translational modifications (PTMs) of serine like phosphorylation or glycosylation can block transfer of the GAG-priming xylose.

Two D-Gal and one D-GlcA units are then added to the xylosylated protein by galactosyltransferases (GalT-I and -II) and β-1,3-glucuronosyltransferase I (GlcAT-I) using UDP-Gal and UDP-GlcA respectively, in order to form the universal GAG attachment tetrasaccharide linker GlcA-Gal-Gal-Xyl. The next added saccharide will then determine the type of GAG chain that will be synthesized. An addition of GlcNAc saccharide will induce the assembly of HS/HP chains, while a GalNAc saccharide will orientate the process towards CS/DS production (Figure 9).

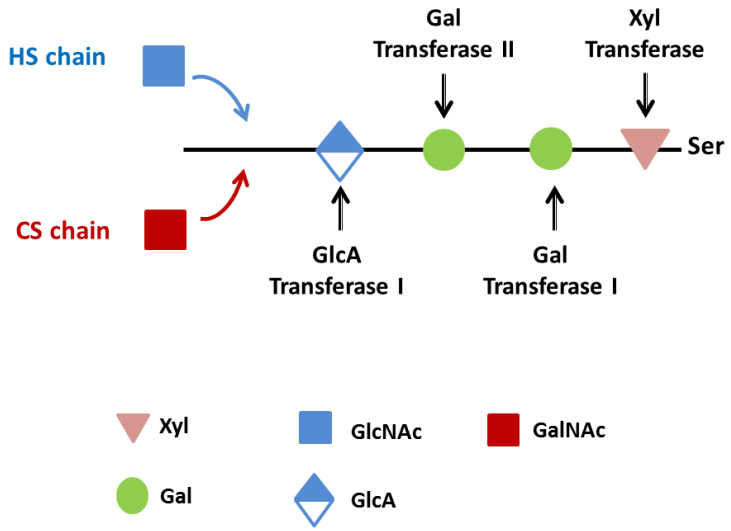


Figure 9: Initiation of HP/HS and CS/DS synthesis. The synthesis of GAGs begins with the transfer on a serine residue of the protein core a xylose, two Gal and a GlcA. The addition of the next saccharide (GlcNAc or GalNAc) will then determine the type of chain.

The GAG biosynthesis is a very complex mechanism involving a large number of actors. The transmission of information from the protein core to the enzymes engaged in the initiation, polymerization and modifications of the GAG chains is still poorly understood.

On one hand, it is not clear whether the GAG chain type is dependent on the protein core. Some PGs are hybrid and one protein core can contain more than one type of GAGs. The choice of GAG synthesized for one specific protein core may be dependent on the cell or the tissue type. For instance, the PG Serglycin contains Heparin in mast cells, but harbors CS chains in other cell types (Prydz, 2015). In addition, the amino-acids residues that surround the Ser-Gly dipeptide can affect the nature of the GAG chain to be synthesized. It has been shown that the acidic and/or hydrophobic (aliphatic and aromatic) residues favor the synthesis of HS chains rather than CS chains (Zhang et al., 1995).

On the other hand, the Xyl can be transitorily phosphorylated to control the formation of GAG chain. The Xyl phosphorylation has been suggested to occur during the initiation process, between the addition of the first Gal and the addition of the GlcA. However, the dephosphorylation of the Xyl seems to be mandatory for the GAG chain synthesis. The Xyl dephosphorylation coincides with the addition of the GlcA by a recently discovered phosphatase XYLP (Koike et al., 2014).

Finally, it has been suggested that the sulfation and phosphorylation of the Gal and GlcA of the tetrasaccharide linker could also influence the type of GAG chain assembled. For example, the sulfation of the Gal is only found in CS (Ueno et al., 2001).

2.2. Elongation of HP/HS

Assembly of the GAG chain is then catalyzed by the alternative transfer of GlcNAc and GlcA residues. It is not known whether the process ends with the presence of GlcNAc or a GlcA. The elongation of HS chains is catalyzed by 5 glycosyltransferases of the extostosin family (EXT1, EXT2, EXTL1, EXTL2 and EXTL3). They are expressed ubiquitously, except for EXTL1, which is highly present in skeletal muscles, the brain and the skin. The EXT proteins are well conserved, especially in their C-terminal region (Busse-Wicher et al., 2014).

The exostosin gene family, when first discovered, were composed of 3 *EXT* genes located in 3 distinct chromosomes, and the first two genes were associated with hereditary multiple osteochondromas (HMO). HMO is an autosomal inherited disorder in which people develop multiple benign bone tumors covered by cartilage. The majority of HMO patients have mutations in *EXT1* (60-70%), occurring throughout the entire length of the gene, and in *EXT2* (30-40%), more specifically located in the N-terminal domain, and resulting all in the expression of truncated EXT proteins (Wuyts and Van Hul, 2000). Unlike EXT3, EXT1 and EXT2 have been cloned and characterized. They were found to share significant homology with each other, but no homology with any other known genes. The knockdown of *EXT1* or *EXT2* in mice results in similar abnormal development (Lin et al., 2000; Stickens et al., 2005). Indeed, they both form a hetero-oligomeric complex capable of polymerizing HS by transferring both GlcA and GlcNAc residues. The activity of only one enzyme is not sufficient for the proper polymerization of the chain and leads to lower glycosyltransferase activity. Moreover, they are not redundant, given that the transfection of EXT2 into EXT1 deficient cells does not restore HS synthesis (McCormick et al., 2000; Wei et al., 2000). It is worth noting that EXT2 may exhibit more functions than the glycosyltransferase activity. For example, it can interact with other proteins implicated in GAG biosynthesis like the maturation enzyme NDST1 (see page 33). This interaction guides NDST1 towards the non-reducing end of the chain where it will achieve its function (Busse et al., 2007; Presto et al., 2008).

More recently, it has been shown that *EXT* enzymes share homology with several *EXT-like* genes that have been identified: *EXTL1, EXTL2* and *EXTL3*. EXTL2 and EXTL3 appear to be

implicated in the HS initiation, given that they can both add the first GlcNAc to the tetrasaccharide linker, and also in HS elongation. EXTL2 may also have additional functions. For instance, it transfers GalNAc, more efficiently than GlcNAc, on the tetrasaccharide linker, suggesting the initiation of a novel glycan (Kitagawa et al., 1999). Moreover, EXTL2 may play a role in the termination of HS elongation. In fact, the addition of GlcNAc by EXTL2 to a linkage region containing phosphorylated Xyl prevents the transfer of more saccharides, stopping thus the HS polymerization process (Figure 10, Nadanaka et al., 2013a). In line with this, it has been shown that a reduced level of EXTL2 in human embryonic kidney (HEK293) cells results in longer HS chains (Katta et al., 2015). In addition, more HS chains were produced in mouse liver cells lacking EXTL2 (Nadanaka et al., 2013b).

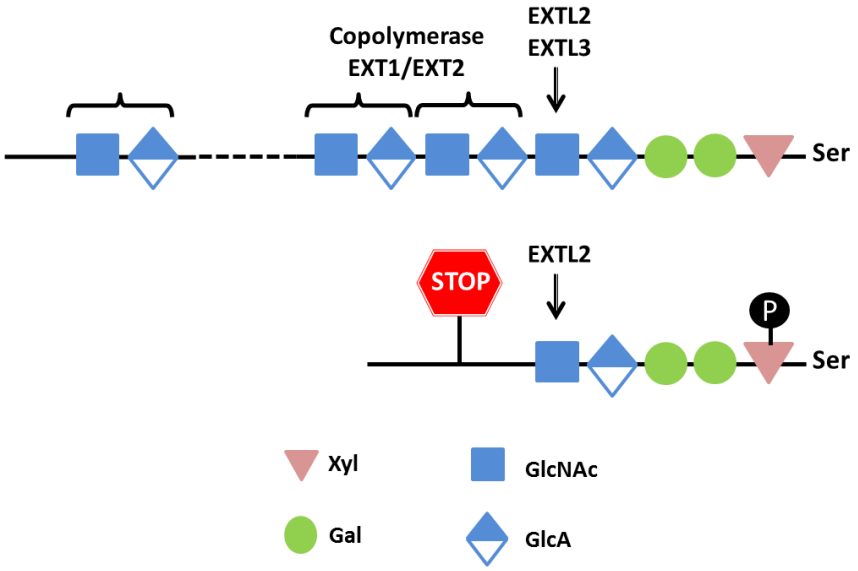


Figure 10: Elongation of HS chain by the action of exostosin. The addition of alternative GlcNAc and GlcA is catalyzed by the complex EXT1/EXT2. EXTL2 and EXTL3 can both add the first GlcNAc saccharides. Finally, EXTL2 can end the elongation of HS by adding a GlcNAc to a tetrasaccharide featuring a phosphorylated xyloside.

Orthologs of EXT1, EXT2 and EXTL3 have been identified in mice, zebrafish, *D. melanogaster* and *Caenorhabditis elegans*. Mutations of *EXT* genes in these animal models lead to developmental abnormalities. This highlights the importance of EXT proteins for HS synthesis and HS interaction with signaling proteins like growth factors and morphogens. Regarding EXTL1 and EXTL2, they are only present in vertebrates (Busse-Wicher et al., 2014). In zebrafish, the EXT1 is expressed by 3 different genes: EXT1a, EXT1b and EXT1c. Interestingly, when zebrafish *EXTL3* is mutated, the HS content is reduced but it is accompanied by an increase in CS content (Holmborn et al., 2012). For *C. elegans*, there are

only 2 EXT family members called Rib1 and Rib2, which are homologous in their amino-acid sequence to mammalian EXT1 and EXTL3, respectively (Kitagawa et al., 2007; Morio et al., 2003). *In vitro*, Rib2, like mammalian EXTL3, exhibits GlcNAc transferase activity. However, the complex of Rib1 and Rib2 is able to transfer both GlcNAc and GlcA sugars. The transfer of GlcA is potentially triggered by Rib1 (Zak et al., 2002).

2.3. Maturation of HP/HS

In the last stage of synthesis, disaccharides undergo controlled modifications by series of enzymatic systems. The first of these reactions is the *N*-deacetylation/*N*-sulfation of glucosamine, catalyzed by enzymes of the NDST family (4 isoforms). This is followed by the epimerization of GlcA into IdoA residues (for HS, HP and DS) by the C5-epimerase and 2-*O*-sulfation of IdoA by the 2-*O*-sulfotransferase (2OST). Finally, sulfate groups at C6 and C3 of the glucosamine are added by the 6-*O*-sulfotransferases (6OST, 3 isoforms) and 3-*O*-sulfotransferase (3OST, 7 isoforms) families, respectively. The sulfation consists in the transfer of the sulfo group from the universal sulfate donator 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to the GAG backbone. PAPS is synthesized in the cytosol and then translocated to the Golgi apparatus by PAPS translocase where it serves as sulfate donor.

Noteworthy, all the enzymes implicated in the biosynthesis of HS (except one 3-*O*-sulfotransferase isozyme) are type II membrane proteins, composed of a short cytoplasmic tail, an hydrophobic transmembrane domain, and a stem region that carries the globular catalytic domain. HS structural features are thus finely tuned by the concerted action of these enzyme families and the differential expression of their multiple isoforms (Esko and Lindahl, 2001; Esko and Selleck, 2002; Kusche-Gullberg and Kjellén, 2003; Li and Kusche-Gullberg, 2016) (Figure 11).

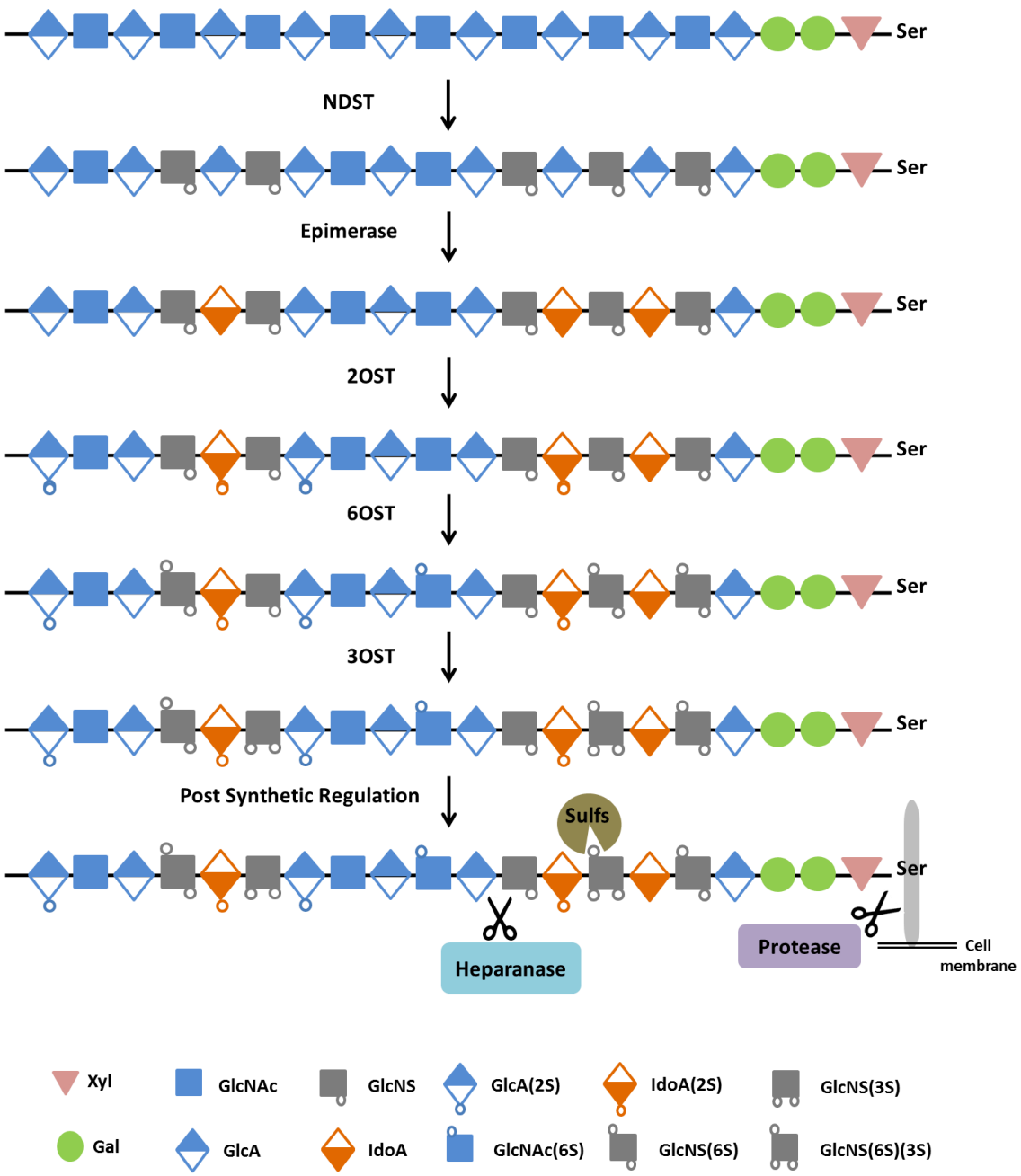


Figure 11: Maturation of HP/HS. Modification steps during HS biosynthesis. Elongating and uniform HP/HS chains successively undergo *N*-deacetylation/*N*-sulfation of glucosamine by the NDSTs, epimerization of GlcA into IdoA by the C5 epimerase, 2-*O*-sulfation of IdoA by the 2OST, 6-*O*-sulfation by the 6OSTs and finally 3-*O*-sulfation by the 3OSTs. The mature HP/HS chains can then be modified post-synthetically by the action of proteases, heparanase and Sulfs.

2.3.1. NDST

The activity of NDSTs is the first step of HS maturation, which has to take place for further modifications to occur, especially the GlcA into IdoA epimerization, 2-*O*- and 3-*O*-sulfation. Inhibition of NDST1 in Chinese-hamster ovary (CHO) cells results in a decrease of

epimerization and *O*-sulfation (Bame et al., 1991a, 1991b, 1994). However, some 6-*O*-S groups can be found in HS unmodified with NDST (Holmborn et al., 2004). NDSTs exhibit two different functions: the removal of the *N*-acetyl group from the GlcNAc, followed by the addition of a sulfate group instead. These two functions are carried out by two distinct catalytic sites within the protein: the *N*-deacetylase activity by N-terminal domain, and *N*-sulfotransferase activity by the C-terminal domain (Berninsone and Hirschberg, 1998). They both work cooperatively for optimal enzyme activity. Indeed, the inhibition of *N*-sulfotransferase activity decreases the rate of *N*-deacetylation (Dou et al., 2015).

NDST is a large family of type II membrane proteins composed of 4 isoforms in mammalian cells, which share high sequence similarity among isoforms and between species: NDST1, 2, 3 and 4 (Aikawa and Esko, 1999; Aikawa et al., 2001). However, they differ in their stages, levels and sites of expression. NDST1 and NDST2 are ubiquitously expressed during embryonic and in adult stage. NDST1 is found in different tissues, while the expression of NDST2 is limited to mast cells. The complete inactivation of NDST1 in mice leads to respiratory distress causing death (Fan et al., 2000). NDST3 and 4 are mostly produced during embryonic development, although NDST3 can also be found in adult brain. Mice that do not produce NDST3 are viable and show only some abnormal behavior (Pallerla et al., 2008). Moreover, levels of expression can differ in pathological conditions like in inflammatory processes. For example, vascular lesions in mice lead to an important increase of NDST1 expression (Adhikari et al., 2008). Another similar study showed that the stimulation of endothelial cells line by pro-inflammatory factors like lipopolysaccharide (LPS) or tumor necrosis factor (TNFα) results in the decrease of NDST2 and NDST3 expression, but induces the expression of NDST1 (Berninsone and Hirschberg, 1998; Krenn et al., 2008).

In addition, NDSTs differ by their substrate specificities. The activity of the 4 enzyme isoforms was studied and showed significant differences. NDST3 exhibits very strong deacetylase activity but weak sulfotransferase activity, and thus has a very high deactylase/sulfotransferase ratio (10.5). Other NDSTs have opposite properties, especially NDST4, which shows a very low deacetylase/sulfotransferase activity ratio (0.04). Ratios of NDST1 and 2 are moderate, with a higher one for NDST2 (Aikawa et al., 2001). This is consistent with the idea that a GlcNAc could be deacetylated by one NDST, but sulfated by another one. These differences in activity could be explained by preferences for different substrates. Indeed, molecular modeling of the sulfotransferase domains of the murine and human NDSTs showed varying surface charge distributions within the substrate binding cleft

(Aikawa et al., 2001). The substrate preference of NDST1 and NDST2 has been extensively investigated. NDST1 acts preferentially on HS that has not been fully sulfated and modified. In fact, NDST1 binds to its substrate in a random way and moves from the non-reducing end towards the reducing end in an oriented process, converting the GlcNAc residue into GlcNS, thereby resulting in the formation of GlcNS clusters. This process ends when NDST1 reaches a premodified GlcNS residue five units away from the reducing end, which results in GlcNAc remaining intact (Figure 12). This study highlighted the important role played by NDST1 in the distribution of NS and NAc domains along the HS chain (Sheng et al., 2011). Regarding NDST2, its overexpression in human embryonic kidney cells generates HS with longer *N*-sulfated regions and higher degree of sulfation, compared to those found in NDST1 overexpressed cells (Pikas et al., 2000). Interestingly, the presence of NDST2 in embryonic liver does not affect HS structure as long as NDST1 is also present. However, in contrast to NDST1, NDST2 targets low sulfated HS to extend further NS domain (Figure 12, Ledin et al., 2006).

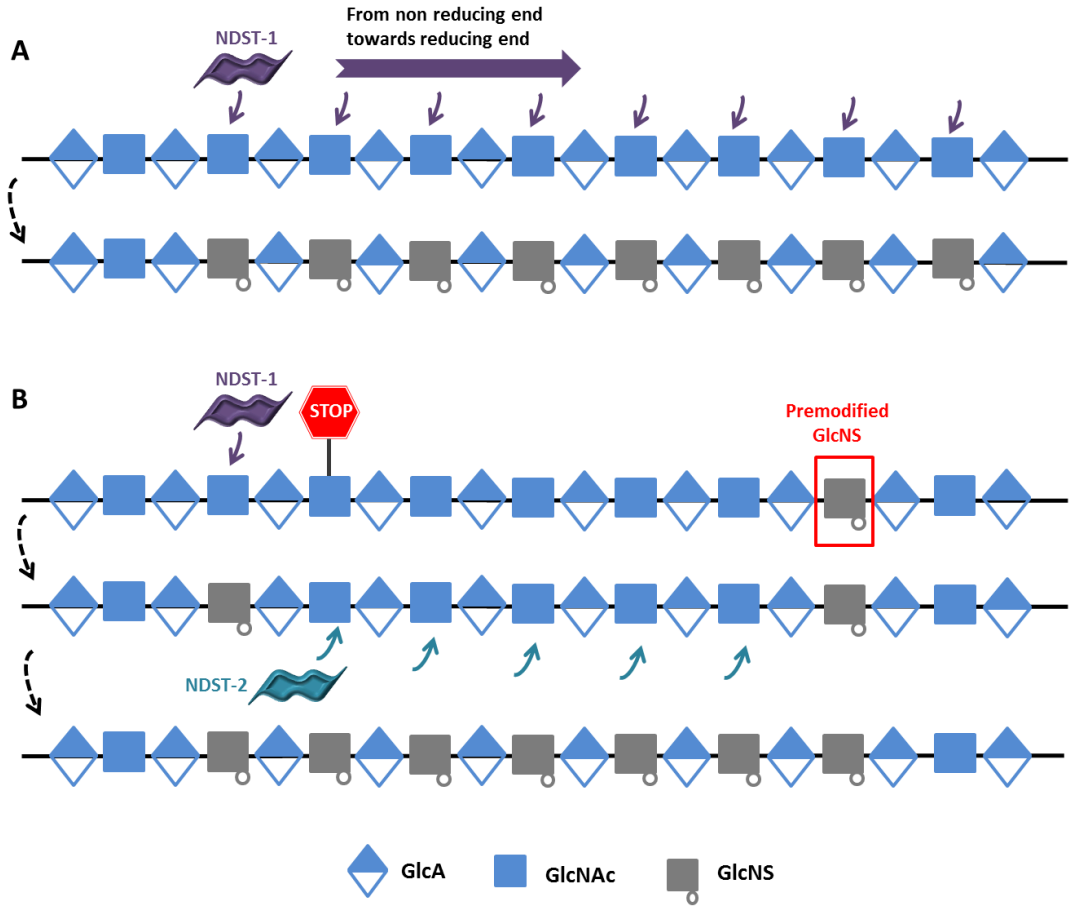


Figure 12: Mode of action of NDST1 and NDST2. NDST1 converts the GlcNAc residue into GlcNS, by moving from the non-reducing end towards the reducing end. This process ends when NDST1 reaches a premodified GlcNS residue five units away from the reducing end. NDST2 can further modify the remaining *N*-acetylated sugars.

To conclude, the expression of the adequate isoform depends on the tissue, the stage of development and the cell needs. In this way, some NDSTs could initiate the *N*-deacetylation/*N*-sulfation while others could extend the stretch of modified residues, in order to adjust the length of the GlcNS. This regulatory system may represent a code that may be read by following biosynthesis enzymes (Zhang et al., 2016).

2.3.2. Epimerase

The glucuronyl C5 epimerase is responsible for the conversion of GlcA to IdoA, increasing thus the flexibility of HS, given that IdoA can adopt different conformations, which facilitate the interaction of HS with its ligands. The epimerization takes place after the NDST activity. This enzyme reaction consists in the removal and the re-addition of the C5 proton *via* a carbanion intermediate, with an inversion of configuration in a way that the carboxyl group is shifted across the plane of the pyranose ring (Figure 13, Lindahl et al., 1976). In a soluble system, the reaction is reversible and attends the equilibrium (Hagner-Mcwhirter et al., 2000). However, in a cellular system, the reaction appears to be irreversible (Hagner-McWhirter et al., 2004) suggesting that further *O*-sulfation promotes the generation of IdoA, by blocking the reversibility of the epimerization. Indeed, IdoA residues are more susceptible to be 2-*O*-sulfated than GlcA residues. In fact, the C5 epimerase can associate with the 2-*O*-sulfotransferase in order to improve its stability, its efficiency and its translocation to the Golgi (Pinhal et al., 2001; Préchoux et al., 2015).

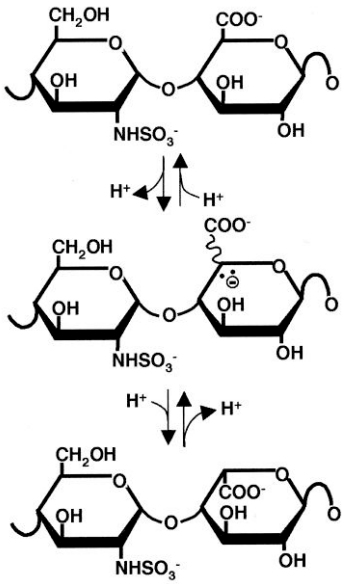


Figure 13: Proposed reaction mechanism for C5 epimerase. C5 epimerization involves abstraction of the C5 proton of GlcA followed by re-addition of a proton from the medium to the resultant carbanion intermediate to generate IdoA. In a soluble system, the reaction is freely reversible. From (Hagner-McWhirter et al., 2004).

The epimerase is a type II transmembrane protein represented by only one gene in the mammalian genome. The crystal structure of zebrafish and human epimerase was achieved alone or in the presence of an oligosaccharide as a ligand and showed a dimeric organization. Each subunit represents a positively charged C-terminal α-helical that are interconnected and that comprise the two catalytic sites, both holding the negatively charged ligand. Three tyrosines appears to be crucial for the enzyme activity (Debarnot et al., 2019; Qin et al., 2015).

The importance of the epimerase has been highlighted in many physiological processes like lymphangiogenesis, heparin biosynthesis by mast cells, neuronal development, B cell maturation (Bülow and Hobert, 2004; Feyerabend et al., 2006; Reijmers et al., 2010, 2011). In addition, the disruption of the GlcA C5 epimerase gene in mouse embryos resulted in the synthesis of abnormal HS lacking IdoA residues, with a decrease in 2-*O*-sulfation, and interestingly an increase in *N*- and 6-*O*-sulfation to compensate that latter decrease. This disruption leads to mouse neonatal death, with defects of kidney, lung, and in skeletal development (Li et al., 2003). This lethal phenotype may be due to the inhibition of the interaction between HS and signaling proteins, given that the IdoA residue appeared to be present in most HS epitopes that bind to proteins.

Moreover, the epimerase gene may act as a tumor suppressor, since downregulation of the enzyme was reported in breast and lung cancers (Grigorieva, 2011; Grigorieva et al., 2008). The epimerization activity is directly correlated with the expression of the C5 epimerase and different systems appear to control this expression in cancer. For example, the transactivation of β-catenin/T-cell factor 4 complex transcriptionally modulates the C5 epimerase expression in human colon cancer cells, and in breast tumors *in vitro* and *in vivo* (Ghiselli and Agrawal, 2005; Mostovich et al., 2012). It has also been suggested that chromatin structure is involved in this regulation (Mostovich et al., 2012). Moreover, micro-RNA miR218 also controls the expression of the epimerase at the post-transcriptional level in colon cancer cells and in breast tumors (Prudnikova et al., 2012; Small et al., 2010).

2.3.3. 2-*O*-sulfotransferase

The 2OST adds sulfate groups in the position two of the uronic acid. It shares some common properties with the C5 epimerase. They are both present as single isoforms and they both have the uronic acid as a substrate. Structural studies have revealed that 2OST recognizes the *N*-sulfate group on its substrate by three essential amino-acid residues (two

arginines and a lysine). However, it does not recognize the 6-*O*-S groups, suggesting that the 2OST activity occurs before the 6OST (Liu et al., 2014). It has also been shown that two histidines are essential for the catalytic activity. The crystal structure of the chicken 2OST showed that the enzyme can form a trimeric complex that appears to be essential for its enzymatic activity, where the three active sites can act independently. Interestingly, some residues implicated in the substrate specificity were identified (Bethea et al., 2008).

The KO of 2OST resulted in neonatal death, renal agenesis and embryonic development retardation in mice (Bullock et al., 1998), and in perturbation of nervous system development in *C. elegans* (Kinnunen et al., 2005). It appears that this KO is less dramatical in *Drosophila*, suggesting that the loss of 2-*O*-S can be compensated by an increase in 6-*O*-sulfation (Kamimura et al., 2006). This compensatory effect was also shown in mouse endothelial cells, where 2OST disruption results in an increase of neutrophil filtration during acute inflammation, due to an enhanced 6-*O*- and *N*-sulfation (Axelsson et al., 2012).

The 2OST is able to transfer sulfate group to both GlcA and IdoA (Rong et al., 2000) but it has a remarkably stronger substrate preference for the IdoA (Figure 14, Rong et al., 2001). Once the 2OST activity takes place, the epimerization is irreversible. Moreover, it has been shown that mono 2-*O*-sulfated octasaccharides are better substrates than octasaccharides with no 2-*O*-S groups (Smeds et al., 2010).

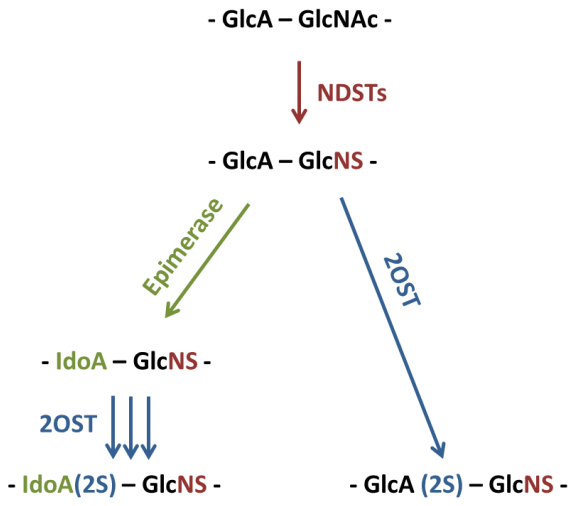


Figure 14: Action of 2OST. 2OST can add sulfate group on the C2 of both GlcA and IdoA, with a preference to IdoA.

Regarding GlcA(2S), it is rare but is found in higher level in adult human cerebral cortex, and in nuclear fractions from hepatocytes (Fedarko and Conrad, 1986; Lindahl et al., 1995).

On the contrary, IdoA(2S) are found in most HS binding protein motifs. It binds to ATIII, FGF2, lipoprotein lipase, HGF and platelet derived growth factor (PDGF) (Jemth et al., 2002; Kreuger et al., 2001; Lyon et al., 1994; Maccarana et al., 1993).

2.3.4. 3-*O*-sulfotransferase

The addition of 3-*O*-sulfation on the GlcNAc residue by the 3OSTs occurs in the last steps of HS maturation and is considered as one of the rarest modifications (one each 100 disaccharides, Kusche et al., 1988). It is present in a limited number of HS chains or can be completely absent. Despite this, 3OST is the largest family of biosynthesis enzymes, gathering 7 isoforms in mammalian species (3OST1, 2, 3A, 3B, 4, 5 and 6) and 8 in zebrafish (Cadwallader and Yost, 2006). These different isoforms share more than 60% of homology in their amino-acid sequences (Shworak et al., 1999). They are all transmembrane proteins except for 3OST1 that lacks a transmembrane region. Its location in the Golgi may thus be due to interactions with other proteins that reside there (Liu et al., 1999; Shworak et al., 1999). The importance of the 3-*O*-S groups was mostly revealed by studies of the interaction between HS and ATIII, a natural anticoagulant factor that regulates the blood coagulation cascade. ATIII binds to specific saccharide motif like [Glc(NS or NAc)(6S), GlcA, GlcNS(3S±6S), IdoA(2S), GlcNS(6S)]. The 3-*O*-S group then triggers ATIII conformational change, which improves its activity. The addition of this 3-*O*-S is specifically catalyzed by the 3OST1 and 3OST5, suggesting the existence of distinct substrate specificities among isoforms (Liu et al., 1999; Shworak et al., 1999; Xia et al., 2002). Indeed, 3OST1 targets preferentially glucosamine saccharides coupled at their non-reducing end to uronic acid devoid of 2-*O*-S groups (Figure 15, Thacker et al., 2014). In contrast, other sulfotransferases like 3OST2, 3OST3a, 3OST3b, 3OST4, 3OST6, and even 3OST5 are involved in the generation of saccharides motifs that serve as an entry receptor for herpes simplex virus 1 (HSV-1) through binding to viral envelope glycoprotein D (Xia et al., 2002). Interestingly, these isoforms prefer the presence of 2-*O*-S groups in the latter position. Indeed, 3OST2 transfers a sulfate group to [GlcA(2S), GlcNS] and [IdoA(2S), GlcNS] motifs, while 3OST3a transfers a sulfate to [IdoA(2S), GlcNS] motifs (Figure 15).

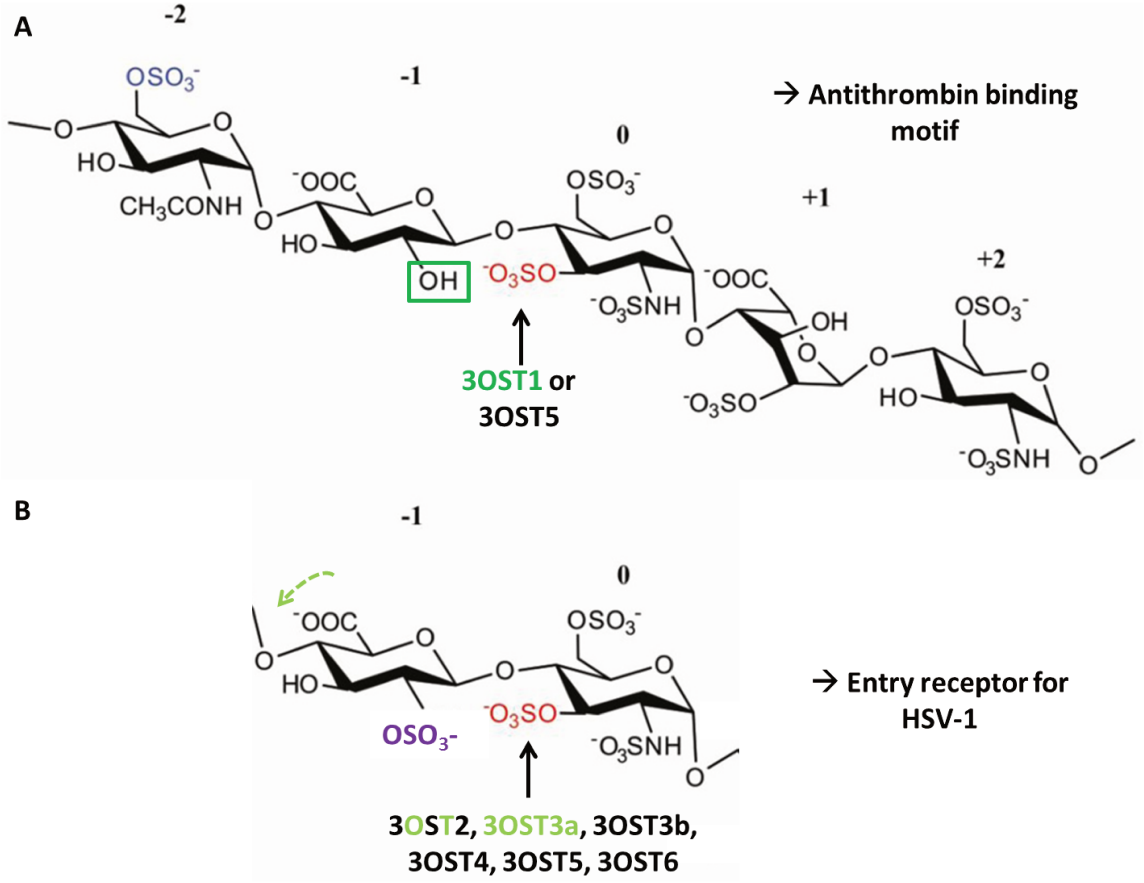


Figure 15: Importance of 3-*O*-S groups in HS/antithrombin and in the entry receptor for HSV-1. The saccharides are numbered relative to the GlcNS(3S) residue at position 0. The binding to antithrombin requires the central 3-*O*-S group (in red) and the 6-*O*-S group at residue −2 (in blue). The 3-*O*-S is added by 3OST1 (only if the -1 residue lacks the 2-*O*-S) or by 3OST5. Note that the 6-*O*-S group on the 3-*O*-sulfated glucosamine residue is dispensable and that residue −2 can be *N*-sulfated (A). The entry receptor for HSV-1 requires the central 3-*O*-S group (in red) and the 2-*O*-S group at residue −1 (in purple). In this case, the 3-*O*-S is generated by 3OST3a (only if the -1 residue is an IdoA), 3OST4, 3OST5, 3OST6 (only if the -1 residue is a GlcA), and finally by 3OST2 (no preference for residue -1). Adapted from (Thacker et al., 2014).

These findings were supported by structural studies that highlighted the mechanism of HS recognition for the different isoforms. Crystal structures of sulfotransferases (3OST1, 3OST3 and 3OST5) in the presence of PAPS and oligosaccharides revealed the presence of two essential domains. First, a catalytic domain, which is highly similar to that of other sulfotransferases, is responsible for binding the acceptor and is essential to the enzyme activity. Catalytic domains of all 3OSTs use a conserved glutamate residue as a catalytic base (Edavettal et al., 2004). The second domain, outside the catalytic domain, is called “the gate” and contains residues involved in substrate specificity. Interestingly, inverting these residues between 3OST1 and 3OST5 is sufficient to convert the substrate specificity of one enzyme to the other (Moon et al., 2004, 2012; Xu et al., 2008).

The expression of 3OSTs differs among isoforms and is controlled in a spatiotemporal manner. 3OST1 and 3OST3a and b are extensively detected in many organs. However, 3OST2 and O3ST4 are expressed in the brain, 3OST5 in skeletal muscle and 3OST6 mostly in the liver and kidney but can be found in small level in the heart, the brain, the lung and the testis (Thacker et al., 2014).

2.3.5. 6-*O*-sulfotransferase

The 6-*O*-sulfation of HS is directed by the 6OSTs, which catalyze the transfer of a sulfate group from PAPS donor to position 6 of glucosamine residues. They exist as 3 isoforms. 6OST1 was first cloned and characterized from CHO in 1995 (Habuchi et al., 1995). On the basis of sequence homology, 6OST2 and 6OST3 isoforms were subsequently identified (Habuchi et al., 2000), as well as an alternatively spliced form of 6OST2, 6OST2-S, featuring a 40 residues deletion (Habuchi et al., 2003). Orthologs are found in Xenopus, *C. elegans* and *Drosophila* (one isoform), while 2 isoforms have been described in chicken, 4 in zebrafish and 3 in human and mouse. In human, 6OST isoforms are encoded by 3 distinct genes located on chromosomes 2, X and 13, respectively (Habuchi et al., 2003). Amino-acid sequences of 6OSTs are less conserved than other glucosaminyl-sulfotransferases. 6OST1 displays 51% and 57% sequence identity with 6OST2 and 6OST3 respectively, while 6OST2 and 6OST3 share 50 % of similarity (Habuchi et al., 2000). However, the sequence located in the central region at the level of potential PAPS binding sites is highly conserved between the three isoforms (Habuchi, 2000). 6OSTs are type II transmembrane proteins that reside in the Golgi apparatus. They feature an N-terminal stem region, which is essential for controlling protein trafficking and localization, for ensuring the oligomer formation and for maintaining the enzyme in an active state (Nagai et al., 2004). Contrary to most other HS biosynthesis enzymes, 6OST can also be found in the extracellular environment (Habuchi et al., 1995). The biological relevance and mechanisms underlying enzyme secretion still remain unclear, but for 6OST3, it has been shown to involve the cleavage of the short hydrophobic cytoplasmic domain by β-secretase (Nagai et al., 2007). Interestingly, inhibition of β-secretase resulted in 6OST3 accumulation in the Golgi apparatus and increased HS 6-*O*-sulfation. The balance of 6OST intracellular/extracellular distribution in active or inactive forms could therefore contribute to the tuning of HS 6-*O*-sulfation. During HS biosynthesis, most steps of polysaccharide assembly and maturation are controlled by enzyme families. Such occurrence of multiple isoforms with specific activities, substrate preferences and tissue distribution is believed to be fundamental for determining the polysaccharide fine structure.

Rather surprisingly, studies of 6OST did not reveal major differences in isoform substrate specificities. All isoforms can indistinctly add sulfate groups on both GlcNAc and GlcNS, regardless of the nature of the adjacent uronic acid (IdoA or GlcA) (Smeds et al., 2003), and overexpression of any of the 6OST in HEK cells affected similarly HS structure (Do et al., 2006). However, studies using libraries of structurally defined oligosaccharides revealed some subtle substrate preferences. 6OST1 preferentially acts on [IdoA, GlcNS] motifs and generate 6-*O*-sulfated sequences with lower 2-*O*-sulfate content. 6OST2 is more active towards 2-*O*-sulfate containing substrates (Jemth et al., 2003). Finally, 6OST3 displays intermediate substrate specificity, between 6OST1 and 6OST2 (Figure 16).

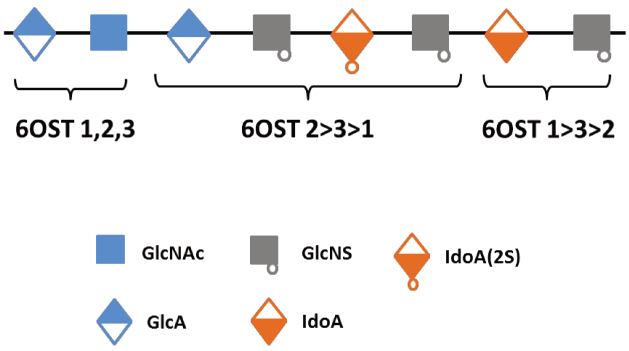


Figure 16: Substrate references of 6OSTs.

In contrast, 6OST isoforms can be distinguished by specific spatio-temporal distribution in tissues (Habuchi et al., 2000, 2003; Sedita et al., 2004). In adult mice, 6OST1 transcripts are mostly present in liver, with moderate expression detected in brain, heart, kidney and lung. 6OST2 is found in the brain and spleen, while its alternatively spliced variant 6OST2-S is more present in the ovary, placenta and fetal kidney. Finally, 6OST3 transcripts are ubiquitously expressed. 6OSTs are also differentially expressed during organogenesis in a stage-dependent manner, suggesting critical roles during development. 6OST1 is found in tissues of epithelial and neuronal origins, while 6OST2, is localized in mesenchymal tissues. In contrast, 6OST3 expression is much more restricted and is limited to the later stages of development. Noteworthy, expression of HS binding morphogens with distinct 6-*O*-sulfation requirements varies extensively during development. Differential expression of 6OST isoforms may therefore provide HS structures with defined binding properties, to finely regulate these morphogens activities and induce specific development processes. Altogether, these data clearly indicate that 6OST tune HS 6-*O*-sulfation through overlapping yet complementary activities.

Regulatory mechanisms involved remain far from being fully apprehended, but interesting data came from studies on animals, in which 6OST have been downregulated or knocked out. In this regard, *Drosophila melanogaster* and *C. elegans* provided valuable models, as these organisms present only one isoform of 6OST. In Drosophila, where structural properties of HS are similar to those present in vertebrates, knockdown of 6OST leads to high lethality and defects in the migration of the tracheal cells, in which 6OST is specifically expressed (Kamimura et al., 2001). Interestingly, these phenotypes are similar to those observed for mutants of FGF signaling pathways, suggesting that these may due to impaired FGF signaling. In *C. elegans*, the unique 6OST form is only expressed in neuronal tissues. Consistently, 6OST KO leads to defects in the ventral cord interneurons (Bülow and Hobert, 2004). In mouse, the KO of 6OST1 leads to high level of lethality during late embryonic stages. Surviving mice are fertile but show growth retards and abnormal morphological phenotypes, such as impaired ossification, reduction of the body weight, defects in placental vascularization, impaired lung morphology and erroneous axon navigation at the optic chiasm (Habuchi et al., 2007; Pratt et al., 2006). In contrast, the KO of 6OST2 did not cause significant phenotype abnormalities, whereas 6OST1/6OST2 double KO mice died at earlier stage than 6OST1 KO mice (Sugaya et al., 2008). Analysis of HS composition from various organs of KO mice showed a reduction in 6-*O*-sulfation content. For 6OST1 KO mice, 6-*O*-sulfation reduction was moderate, with more pronounced effects in tissues naturally expressing high levels of this isoform. Transcriptional analysis showed no increase in 6OST2 and 6OST3, thereby indicating an absence of isoform compensatory mechanisms in these mice. Noteworthy, [IdoA(2S), GlcNS(6S)] units were less affected than other 6-*O*-sulfate containing disaccharides. This suggests that 6OST1 may not be primarily involved in the 6-*O*-sulfation of heparin and highly sulfated HS. In contrast, analysis of HS from 6OST2 KO mouse embryonic fibroblasts (MEFs) showed a marked decrease in 6-*O*-sulfation, [IdoA(2S), GlcNS(6S)] being the disaccharide unit the most affected. Finally, 6OST1/6OST2 double KO resulted in an almost complete loss of 6-*O*-sulfation. Interestingly, analysis of double KO MEF showed an increase in 2-*O*-sulfates, suggesting a possible compensatory effect between OSTs. In line with this, overexpression of any of the 6OST led to an increase in 6-*O*-sulfation, accompanied by a decrease in 2-*O*-sulfation (Do et al., 2006). Such compensation mechanism is still unclear and may extensively vary amongst tissues and species, but this clearly underlines the tight connections between the various enzymatic modification steps during HS biosynthesis.

2.4. Concept of Gagosome

It is now clearly believed that the biosynthesis enzymes, all located in the Golgi compartment, are organized into large complexes, called gagosome, that cooperatively and finely control the HS elongation and maturation, depending on the cells' need, in order to generate a large diversity of HS structural motifs with precise functions. Investigations of this concept have been so far restricted to studying interactions of HS biosynthesis enzymes in pair and not in bigger complexes (Figure 17). The hypothesis of the gagosome existence arose with the observation that overexpression of NDST1 and NDST2 in HEK293 cells differently affected HS sulfation, as mentioned before, suggesting their interaction with other maturation enzymes (Ledin et al., 2006; Pikas et al., 2000). Subsequently, it was shown that NDST1 binds to EXT2. This binding may protect NDST1 from degradation or guide it to its site of action in the Golgi. In addition, the overexpression of EXT2 triggers the expression of NDST1. However, amounts of NDST1 is reduced when EXT1 is overexpressed, suggesting that NDST1 and EXT1 compete for binding to EXT2, and unbound NDST are degraded in the ER (Presto et al., 2008). Indeed, EXT1 and EXT2 form together an hetero-oligomeric complex accumulating in the Golgi, which has a higher glycosyltransferase activity than EXT1 alone (McCormick et al., 2000). However, it is not clear whether the NDST1 and EXT2 interaction occurs in the ER or the Golgi or if it is already established in the ER and persists within the Golgi.

The presence of the gagosome reseau was also suggested thanks to the importance of IdoA(2S) residues that guided the research to clarify the link between the addition of 2-*O*-S groups and the epimerization and to understand whether these modifications were random or concerted. In that context, it has been shown that 2OST and C5 epimerase interact together physically with high affinity, to form a complex that is required for the epimerase stability, activity of the enzymes, translocation to the Golgi, and for generating processively more extended domains of [GlcNS, IdoA(2S)] repeats. This association may thus ensure the rapidity of HS maturation (Pinhal et al., 2001; Préchoux et al., 2015).

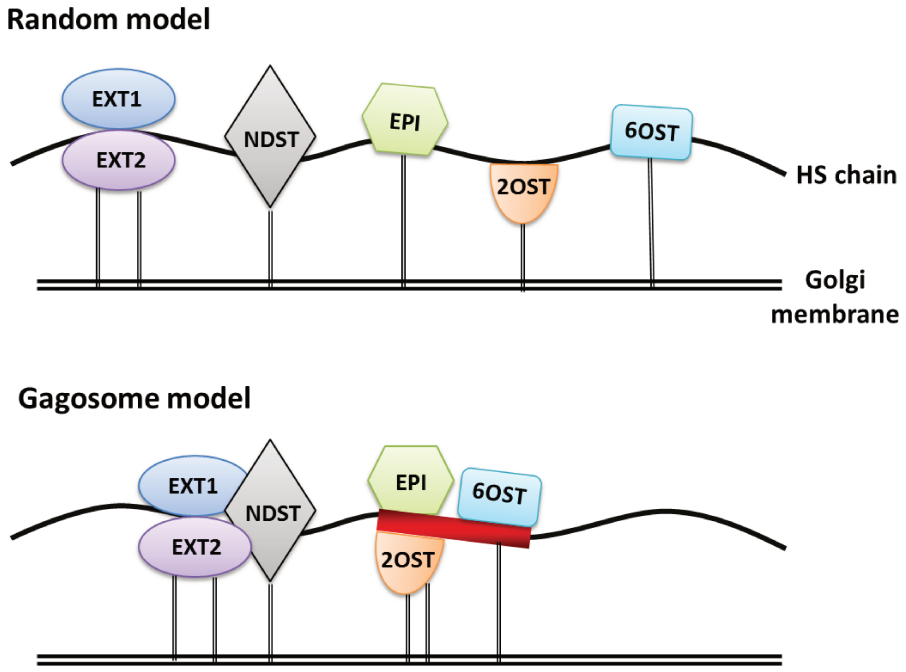


Figure 17: Gagosome model

2.5. Elongation and maturation of CS

After the formation of the tetrasaccharide linker, a first GalNAc is added by the N-acetylglactosyl transferase I, orientating the fate of GAG chain towards CS. Next, the alternative transfer of GalNAc and GlcA, by N-acetylglactosyl transferase II and glucuronic acid transferase II respectively, catalyzes the polymerization of the CS chain. An additional enzyme capable to add both GalNAc and GlcA, called chondroitin synthase, was identified. The elongation of CS chain by this enzyme requires its association to chondroitin polymerizing factor (Mikami and Kitagawa, 2013; Silbert and Sugumaran, 2002).

As mentioned before, GalNAc can be 6-*O*-sulfated, resulting in CS-C disaccharide units. This sulfation is catalyzed by chondroitin-6-*O*-sulfate sulfotransferases (C6STs) that differ in their substrate specificity. C6ST1 is the first characterized and prefers GlcA-rich regions. Noteworthy, it can also add sulfate group to the KS galactose residues (Habuchi et al., 1996). Other homologs of C6ST1 have been identified later. They are involved in 6-*O*-sulfation of residues adjacent to GlcA(2S), given that the KO of that enzyme in mice only affects the level of D units (Uchimura et al., 2002). Another homologous enzyme, referred as GalNAc4S-6ST, targets GalNAc(4S) and generates the highly sulfated CS-E. The 6-*O*-sulfation occurs in the medial/trans Golgi, while 4-*O*-sulfation occurs in a later trans Golgi region.

The 4-*O*-sulfation of GalNAc is catalyzed by 4-*O*-sulfotransferase (C4STs). C4ST1, C4ST2 and C4ST3 prefer GlcA rich regions. Interestingly, C4ST1 cooperates with GalNAcT-II in order to promote CS chain elongation (Mikami and Kitagawa, 2013). D4ST targets the GalNAc adjacent to IdoA (Silbert and Sugumaran, 2002). Once the latter 4-*O*-sulfation takes place, the epimerization becomes irreversible (Mikami and Kitagawa, 2013).

The 2-*O*-sulfation is the last sulfation step in the maturation of CS and DS. It is catalyzed by uronyl 2-O-sulfotransferase (UST). UST can sulfate both GlcA and IdoA, but with a preference for IdoA (Mikami and Kitagawa, 2013; Silbert and Sugumaran, 2002).

Finally, the epimerization of GlcA to IdoA by 2 C5 epimerases DS-epi1 and DS-epi2, results in the formation of DS. IdoA can be either found alterning with GlcA, or clustered together (Pacheco et al., 2009). Interestingly, the epimerization occurs at the same time as the 4-*O*-sulfation (Silbert and Sugumaran, 2002).

The CS chain, once synthesized, can be terminated by either the GlcA or (un)modified GalNAc. However, it has been suggested that the CS-E motif can be a termination signal, given that it is more abundant in the non-reducing end of the CS chains (Midura et al., 1995). In addition, GalNAc4S-6ST KO mice displays larger CS chains (Ohtake-Niimi et al., 2010).

3. Post-synthesis regulation of HP/HS

Apart from biosynthesis, further regulation of HS occurs post-synthetically, through the action of sheddases, extracellular heparanases and extracellular 6-*O*-sulfatases of the Sulfs family (see page 62, Hammond et al., 2014; Rosen and Lemjabbar-Alaoui, 2010; Vivès et al., 2014). All these modifications generate a huge and fine structural and functional diversity of the final product of HS (Figure 11).

Sheddases are enzymes that target the core protein of HSPG like Syndecans, releasing thus soluble HS-containing protein fragments. This process is an important regulation mechanism of the amount of HSPGs found at the cell surface or in the ECM. It can be induced by inflammatory cytokines by triggering intracellular signaling and activating metalloproteinases. Many studies showed that there is correlation between the presence of soluble Syndecan-1 and cancer growth, and that shedding process in myeloma tumors is enhanced by heparanase (Yang et al., 2002, 2007). In addition, it was shown that the shedding results in the removal of sequestered chemokines, thus facilitating the resolution of neutrophilic inflammation (Hayashida et al., 2009). Recently, given that Syndecan-1 was found in the nucleus of myeloma and mesothelioma cells (Chen and Sanderson, 2009; Zong et al., 2009), it has been suggested that shedded Syndecan-1 possesses an unidentified receptor enabling its transport towards the nucleus where it can play specific functions. For example, it has been shown that HS in the nucleus can change the activity of DNA topoisomerase I and histone acetyl transferase HAT (Buczek-Thomas et al., 2008; Kovalszky et al., 1998). Noteworthy, this translocation process has also been shown to be controlled by heparanases (Chen and Sanderson, 2009).

Heparanase belongs to the glycoside hydrolase (GH)79 family that can cleave long HS chains into shorter fragments of 10-20 sugar units. They target the [GlcA, GlcNS] linkages of HS. They prefer the IdoA(2S) for efficient endogenous activity (Bai et al., 1997). Crystal structure of heparanase in the presence of HS showed that heparanase recognizes specifically a trisaccharide that contains *N*-S on the -2 position and a 6-*O*-S groups on the +1 position (Wu et al., 2015).

Heparanase plays both extracellular regulation and intracellular catabolism roles. It can digest extracellular HSPGs or membrane bound HSPGs, releasing HS ligands like growth factors, chemokines and morphogens, and resulting thus in cell proliferation, motility and modeling at inflammation sites. In addition, released HS fragments can also activate

downstream signaling cascades. Heparanase is thus involved in many processes including cell communication, autophagy... (Vlodavsky et al., 2018). Discrete dysregulation of heparanase expression or function can significantly affect the signaling network and cause uncontrolled cell growth, invasion, and activation of immune system… For example, overexpression of heparanase is correlated to cancer metastasis and chemoresistance. In addition, many studies showed that heparanase is associated with various pathologies other than cancer, like inflammation, thrombosis, atherosclerosis, fibrosis, diabetes and kidney disease (Vlodavsky et al., 2018).

Inside the cell, the importance of heparanase is revealed by the fast turnover of HSPGs. Once produced, heparanase is first targeted towards the ER, thanks to its signal peptide, then sent to the Golgi apparatus where it is secreted in a latent form by vesicles that bud from the Golgi. There, it can directly bind to membrane HSPGs and induce their internalization by endocytosis. PG protein core is then digested by proteolysis and the HS chain by glycosyl hydrolases and sulfatases. To facilitate the process, the endosomes can convert to lysosomes, where the heparanase becomes active and process HS chains to generate additional non-reducing ends for the activity of lysosomal exoglycosidases (Vlodavsky et al., 2018). Furthermore, lysosomal heparanase can enter the nucleus where it can interact with the chromatin complex, regulating some histone methylation and gene transcription (Figure 18, He et al., 2012).

Interestingly, heparanase II (Hpa2), an homolog of heparinase has been identified, but lacks enzyme activity, despite retaining conserved critical catalytic residues (McKenzie et al., 2000). However, Hpa2 is able to bind to HS/HP, and even with higher affinity than heparanase, but without activating HS internalization. Hpa2 may thus compete with heparanase for HS binding and inhibit its endoglycosidase activity. It can also inhibit the activity of heparanase by binding directly to it (Levy-Adam et al., 2010). It can therefore promote normal differentiation, apoptosis, and act as a tumor suppressor (Figure 18).

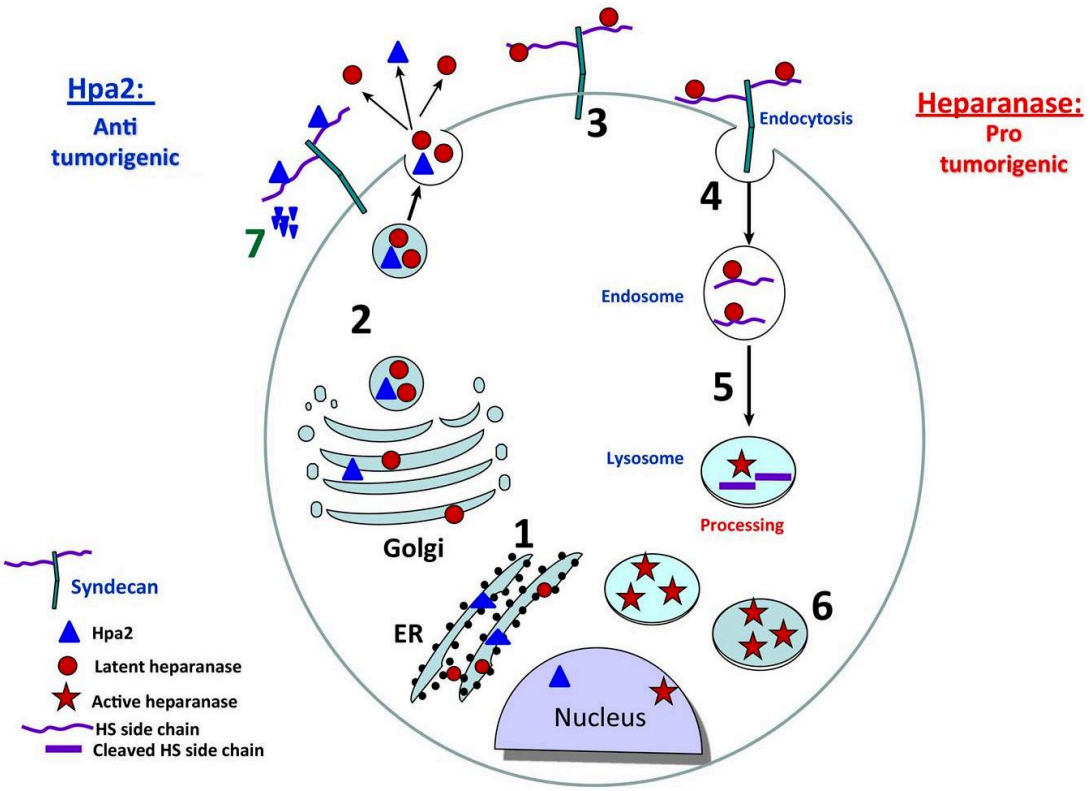


Figure 18: Biosynthesis of heparanase and heparanase II (Hpa2) and trafficking. Latent heparanase is first targeted to the ER lumen by its own signal peptide (1). It is then shuttled to the Golgi apparatus, and is subsequently secreted *via* vesicles that bud from the Golgi (2). Once secreted, heparanase rapidly interacts with syndecans (3), followed by rapid endocytosis of the heparanase/syndecan complexes that accumulate in late endosomes (4). Conversion of endosomes to lysosomes results in heparanase processing and activation (5). Typically, heparanase appears in perinuclear lysosomes (6). Lysosomal heparanase may translocate to the nucleus. Similar to heparanase, Hpa2 is first targeted to the ER lumen (1), secreted *via* vesicles that bud from the Golgi (2) and interacts with syndecan on the cell surface (7). Unlike heparanase, Hpa2 is retained on the cell surface for a relatively long period followed by a decline at later time points, possibly due to proteolysis, or release from the cell surface by shedding of syndecan. Adapted from (Vlodavsky et al., 2018).

According to previous data, heparanase can represent a potential therapeutic target in the cited diseases. Indeed, it has been shown that the downregulation of heparanase can normalize vascular structure in solid tumors, preventing the chemoresistance and promoting the efficient delivery of chemotherapeutic agents (Zhang et al., 2018). In addition, anti-cancer drugs that inhibit heparanase have been developed. For example, PG545 (Pixatimod) are polyanionic molecules that mimic HS and induce apoptosis of lymphoma cells (Weissmann et al., 2018). Monoclonal antibodies have been developed as well, and characterized for their ability to neutralize heparanase and impair lymphoma tumor growth and metastasis (Weissmann et al., 2016).

4. HS/protein interactions – Importance of 6-S groups

Thanks to this finely tuned regulation of HS by biosynthesis and post-synthesis enzymes, HS chains display unique structural features and domain organization and are ready to orchestrate various proteins to play together in order to achieve specific functions (Gallagher, 2015). It is well established that HS/protein interactions involve saccharide motifs with defined sulfation patterns located mostly in the NS domains enriched in trisulfated disaccharide [IdoA(2S), GlcNS(6S)]. It is rare that a protein binds to a single disaccharide unit (Esko 2007). NS domains are distinguished by the conformation they adopt, forming rigid twofold helical symmetry, where the three sulfate groups are on opposite faces of the helical axis. This conformation enables proteins to bind to both sides of the saccharide chain. In addition, the plasticity of the iduronate ring of IdoA(2S) can modify the spatial location of the carboxyl and 2-*O*-S group, results in various structural features with different specificities and affinities for HS ligands (Gallagher, 2015). In addition, NAc domains are flexible domains that could also play important roles in these interactions. In some cases of dimer ligands like IL-8 (CXCL8) and Interferon (IFN), the interaction involves discontinuous domains where the HS chain features two NS domains separated by a NAc domain type linker. In this way, HS can bridge the binding sites located on opposite sides of the dimer (Figure 19, Lortat-Jacob et al., 2002; Sarrazin et al., 2011).

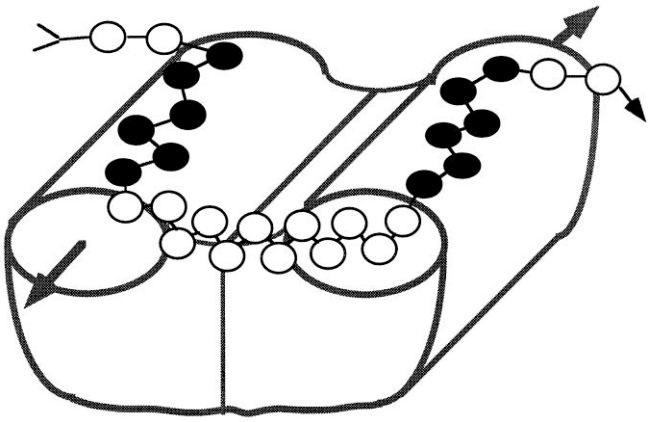


Figure 19: Model of HP/HS binding to the IL-8 dimer. Two identical binding motifs within *N*-sulfated stretches of HS have been shown to interact with the α-helical HP/HS binding domains of each IL-8 monomer. The NS domains (≤ 6 sugar units; closed circles) are bridged by a sequence (≤12–14 monosaccharide units; open circles) that may be either *N*-acetylated or *N*-sulfated. Note the polarity of the sugar chain (arrow). From (Spillmann et al., 1998).

Binding to HS mainly involves electrostatic interactions between positively charged residues of the basic amino-acids on proteins and sulfate and carboxyl groups of the

polysaccharide. However, some acidic chemokines such as CCL3 and CCL4 bind to HS, meaning that the binding is not necessarily based on overall charge interactions (Proudfoot, 2006). The interactions also involve hydrogen bonds and Van der Waals forces.

Regarding the ligand, there are two consensus sequences that can be implicated in heparin binding, identified by Cardin and Weintraub in 1989 and named after them (Cardin and Weintraub, 1989). These patterns are X-B-B-X-B-X and X-B-B-B-X-X-B-X, where B is a basic residue (arginine or lysine) and X is a hydrophobic one. This study showed that, in addition to the overall charge, an arrangement of the basic residues for electrostatic compatibility to HS is necessary. The binding sites can either be presented in a peptide loop of secondary structure, like for FGF and ATIII, or can be located in unstructured regions, as seen in the case of IFN, VEGF and PDGF (Gallagher, 2015). In addition, clusters of more than 4 arginines do not necessarily have higher affinity to HS, binding being even increased in presence of amino-acids residues interlaced between the basic ones (Hileman et al., 1998). It appears that high affinity peptides for HS/HP interaction comprises high levels of arginine or lysine (less), but not histidine, and can also contain important levels of polar amino-acid serine. Regarding hydrogen bounds, aspargine and glutamine enriched peptides have the highest affinity, as shown for FGFs, and tyrosine plays an important role as well in binding to *N*-acetyl group of HP in the case of ATIII (Hileman et al., 1998).

A wealth of studies has reported that expression and structure of HS are highly dynamic and vary dramatically amongst cell types and physiopathological status. This is particularly significant for HS 6-*O*-sulfation pattern, which has been shown to undergo significant changes during both embryonic development (Allen and Rapraeger, 2003; Brickman et al., 1998) and aging (Feyzi et al., 1998; Huynh et al., 2012), as well as in pathologies, such as cancer (Jayson et al., 1998; Safaiyan et al., 1998), amyloidosis (Bruinsma et al., 2010; Hosono-Fukao et al., 2012), chronic renal fibrosis (Alhasan et al., 2014), inflammation (Reine et al., 2012; Wang et al., 2002) and diabetes (Hassing et al., 2012; Wijnhoven et al., 2006). Functionally, HS 6-*O*-sulfation has been associated with major biological functions of the polysaccharide (Figure 20).

Ligand	Function
Antithrombin III	Inhibition of blood coagulation
FGF	Angiogenesis, wound healing, embryonic development, cancer
PDGF	Angiogenesis, wound healing, embryonic development, cancer
HGF	Angiogenesis, tissue regeneration, wound healing, embryonic development, cancer
GDNF	Neuron protection and regeneration, kidney development, spermatogenesis
HB-EGF	Angiogenesis, wound healing, cancer
VEGF	Angiogenesis, cancer
CXCL12	Chemotaxis, inflammation, embryonic development, cancer
CXCL8	Chemotaxis, embryonic development, cancer
CXCL4	Chemotaxis, inflammation, wound healing
CCL5	Chemotaxis, inflammation
TGFβ	Tissue regeneration, embryonic development, cancer
Wnt	Tissue regeneration, embryonic development, cancer
L-selectin	Leukocyte extravasation, inflammation
Endostatin	Inhibition of angiogenesis

Figure 20: Table representing the implication of 6-*O*-S groups in HS/protein interactions.

4.1. HS and anti-thrombin

One of the most studied example of HS/protein interaction is that of the ATIII, which characterization provided the molecular basis for HP anticoagulant properties (Petitou et al., 2003). In fact, HP inhibits and prevents the formation of clots and thrombosis, due to its interaction between with ATIII. That binding causes the activation ATIII by changing its conformation, leading to the inhibition of thrombin, factor Xa and other members of the coagulation cascade. The high affinity binding is mediated through a pentasaccharide motif (see page 39). Critical structural elements of this highly specific saccharide motif are the well-known internal 3-*O*-sulfated glucosamine, but also a 6-*O*-sulfate located on the non-reducing end residue (Zhang et al., 2001).

4.2. HS and growth factors

The selective contribution of 6-*O*-sulfates in HS activities has been most clearly exemplified in the FGF/HS interaction model (Gallagher, 2015; Harmer, 2006; Li et al., 2016; Pye et al., 1998). FGFs represent a family of 23 heparin-binding growth factors. Amongst these, FGF1 and FGF2 are the archetypal members, which provided the first evidence that HS could serve as growth factor co-receptors, which was necessary for promoting their activity (Rapraeger et al., 1991; Yayon et al., 1991). Binding sites to FGF1 and FGF2 share very

similar structural features, both in term of size (5–6 sugar units) and saccharide composition, with a requirement for GlcNS and IdoA(2S) residues (DiGabriele et al., 1998; Faham and Hileman, 1996; Maccarana et al., 1993). However, while 6-*O*-S are essential for binding to FGF1 and enable optimal contact between sugar and protein (Ashikari-Hada et al., 2004; Ishihara, 1994; Kreuger et al., 2001), they are not involved in the interaction with FGF2 (Habuchi et al., 1992; Turnbull et al., 1992). In contrast, induction of FGF2 activity requires longer oligosaccharides (10–12 sugar units) and the presence of 6-*O*-S (Lundin et al., 2000; Pye et al., 1998; Sugaya et al., 2008). These findings provided the first evidence of possible uncoupling between HS interactive properties and biological functions, and highlighted 6-*O*-sulfation as the critical determinant for discriminating these two activities. The rationale behind such mechanism is that saccharide extension and 6-*O*-S provide an additional binding site for FGF receptor (FGFR) that induces formation of FGF/FGFR/HS ternary complex able to trigger cell signaling. Interestingly, it has been recently reported that heparin oligosaccharides featuring a few (1–2) 6-*O*-S located on the reducing end glucosamine residues exhibited full FGF2 promoting activity (Seffouh et al., 2013). The presence as well as the specific positioning of 6-*O*-S within saccharide sequences may therefore be critical for activation of FGF2.

Aside the well-documented examples of FGF1 and FGF2, HS 6-*O*-sulfation has been involved (to various extent) in the binding and/or activation of other FGFs, such as FGF4, 7, 9, 10 and 18 (Ashikari-Hada et al., 2004; Ishihara, 1994; Patel et al., 2008; Sugaya et al., 2008; Xu et al., 2012), as well as other growth factors, including PDGF (Feyzi et al., 1997), HGF (Lyon et al., 1994), the Glial cell line Derived Neurotrophic Factor (GDNF) (Ai et al., 2007; Rickard et al., 2003), Heparin binding Epidermal Growth Factor (HB-EGF) (Cole et al., 2014) and VEGF (Ferreras et al., 2012; Ono et al., 1999; Robinson et al., 2006).

4.3. HS and chemokines, morphogens

HS 6-*O*-sulfation also participates in the interaction with other signaling proteins, such as chemokines and morphogens. Chemokines are small proteins implicated in many biological processes such as development, inflammation and immunosurveillance (Zlotnik and Yoshie, 2012). All chemokines bind to HS, which regulates their activity in different ways. HS sequesters chemokines and protects them from enzymatic cleavage, it increases their local concentration in the ECM to form chemokine gradients guiding migrating cells, and it induces their oligomerization to facilitate their interaction with their receptors (Lortat-Jacob et al., 2002; Monneau et al., 2016; Proudfoot, 2006; Sadir et al., 2004; Sweeney, 2002).

Contribution of 6-*O*-S has been demonstrated for CXCL12 (Roy et al., 2014b; Sadir et al., 2001; Uchimura et al., 2006; Zhang et al., 2012b), CXCL8 (Pichert et al., 2012; Spillmann et al., 1998) and CXCL4 (Platelet factor 4) (Pempe et al., 2012; Stringer and Gallagher, 1997). Interestingly, a recent study showed that addition of a single 6-*O*-sulfate group at the non-reducing end of a chemically synthesized [IdoA(2S), GlcNS]$_6$ heparin oligosaccharide switched its inhibitory properties from CXCL8 toward CXCL12, thereby highlighting the importance of 6-*O*-S positioning for binding to these chemokines (Jayson et al., 2015). Finally, the crystal structure of CCL5 in complex with heparin disaccharides revealed electrostatic interactions between the 6-*O*-S of the sugar and the K_{45} residue of the chemokine (Shaw et al., 2004).

Morphogens are signaling proteins that dictate cell fate and tissue development during embryogenesis. HS binds to members of the 3 major families of mammalian morphogens: Wnt/β-catenin, HH and transforming growth factor beta (TGF-β)/bone morphogenic protein (BMP). As for growth factors and chemokines, interaction with HS regulates morphogen distribution and contributes to the formation of gradients (Coulson-Thomas, 2016; Yan and Lin, 2009). The 6-*O*-sulfation of HS has been reported to modulate activity of both TGFβ and Wnt in two opposite ways. Interaction with HS promotes TGFβ1 activity and 6-*O*-S groups are important structural determinants for the binding (Lyon et al., 1997; Yue et al., 2008). This is in agreement with a recent study showing that a decrease in HS 6-*O*-sulfation reduced the response of primary fibroblasts to TGFβ1 (Lu et al., 2014; Yue et al., 2013). In contrast, 6-*O*-S may act as a negative regulator of Wnt signaling. A proposed mechanism is that Wnt binds with high affinity to 6-*O*-sulfated HS, which prevents access to its cell surface receptor Frizzled (Fz). In support to this, it has been shown that enzymatic removal of 6-*O*-S by the Sulfs (see page 70) reduced Wnt/HS binding affinity, thereby enabling interaction with Fz and induction of cell response (Ai et al., 2003). However, it has also been reported in other studies that 6-*O*-desulfation could also have an inhibitory effect on Wnt, by facilitating its release and degradation (Kleinschmit et al., 2010). Regulation of Wnt signaling may thus be dictated by a complex interplay between HS 6-*O*-sulfation status and Fz bioavailability.

4.4. HS and other ligands

6-*O*-S have also been found to participate to the binding of HS to L-selectin (Wang et al., 2002; Zhang et al., 2012b), endostatin (Blackhall et al., 2003) and axon guidance protein slit-2 (Shipp and Hsieh-Wilson, 2007). It has also been involved in cell surface attachment of hepatitis E virus (Kalia et al., 2009) and in the promotion of neuregulin-1/erbB receptor

interaction (Pankonin et al., 2005). Finally, a number of studies have shown evidence of an implication in Alzheimer, as 6-*O*-S take part in the binding to β-amyloid peptides (Lindahl et al., 1999) and in the modulation of amyloid precursor protein processing (Scholefield et al., 2003).

5. 6-*O*-desulfation of HS by Sulfs

In addition to the biosynthesis processes *via* the 6OSTs, HS 6-*O*-sulfation is further regulated through a post-synthesis mechanism involving extracellular sulfatases of the unique Sulfs family. Sulfs result in little modifications of the structure of HS, but by targeting specifically the 6-*O*-S groups, which are involved in the binding of many signaling proteins, they cause great functional consequences.

5.1. Generalities on sulfatases

The sulfatase family is a group of enzymes that catalyze the hydrolysis of sulfate ester bonds from a large array of sulfated substrates such as steroids, glycolipids, and proteoglycans. They remove as well sulfates from sulfamate groups (C-*N*-S). They are implicated in many physiological processes like hormone regulation, cellular degradation and the control of signaling pathways. Regardless their substrates, they share homologies in their sequence (20-60%), their structure and their activity and they are conserved among prokaryotic and eukaryotic species. They all feature, in their N-terminal region, two highly conserved signature sequences that belong to the active site of the enzyme.

The first one is a five amino-acid peptide **C/S-X-P-S/X-R** that starts with a cysteine converted post-translationally to a formylglycine (FGly) (Dierks et al., 1999; Knaust et al., 1998). FGly is the only naturally occurring amino-acid residue that has an aldehyde functional group and that is essential for the enzymatic activity. However, for some prokaryotic species, the FGly results from the oxidation of a serine instead of a cysteine (Dierks et al., 1998a; Miech et al., 1998). The proline and the arginine play important role in the direction of the FGly modification and in the structural organization of the active site. The second signature is a sequence **G-K-X-X-H** where the lysine and the histidine are important for the sulfate ester catalysis (Waldow et al., 1999). The active site peptide includes also a divalent metal ion located within a pocket in which substrates are bound.

Regarding the sulfatase activity, two mechanisms have been suggested (Hanson et al., 2004) (Figure 21). Either the FGly residue serves directly as an electrophile forming a sulfate diester *via* an addition-hydrolysis mechanism, or the FGly acts as an aldehyde-hydrate by a transesterification elimination mechanism (Hanson et al., 2004). This results in the release of an alcoholate, but the mechanism of release of the bound sulfate is not clear yet (Marino et al. 2013).

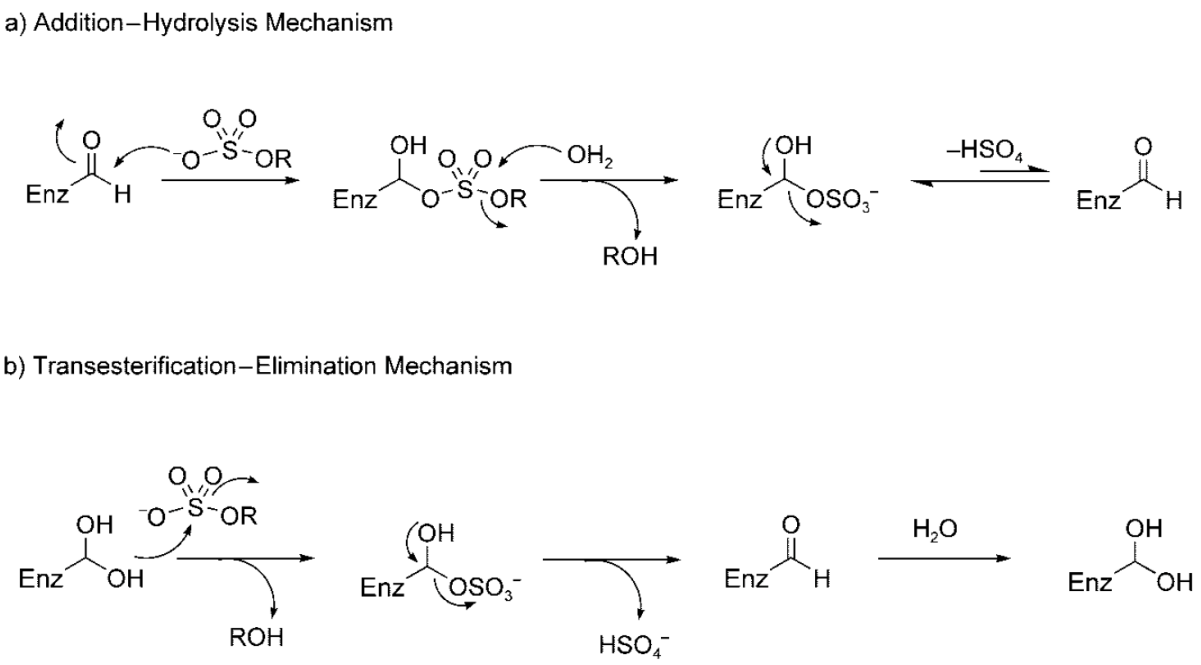


Figure 21: Proposed mechanistic schemes for the hydrolysis of sulfate esters by the active site aldehyde FGly. From (Hanson et al., 2004).

In eukaryotes, the cysteine conversion is catalyzed by a formylglycine generating enzyme encoded by sulfatase modifying factor 1 (SUMF1). The importance of the desulfation processes is highlighted by the number of diseases that result from the uncontrolled accumulation of sulfated compounds. For example, multiple sulfatase deficiency (MSD) is a rare human disorder caused by a mutation of *SUMF1* gene that prevents the FGly formation and results in defective activity of all sulfatases (Schmidt et al., 1995). A paralogue of SUMF1, SUMF2 has also been identified by sequence homology (Cosma et al., 2003). It exhibits the same activity although with less efficiency than SUMF1, and it has thus been suggested to be the responsible for the little sulfatase activity found in MSD.

Homologous genes of SUMF1 were also detected in bacteria by sequence analysis. The cysteine FGly conversion enzymes appeared to be more active in prokaryotes than in eukaryotes. All the cysteines are converted into FGly in prokaryotic sulfatases (PARS). However, not all the cysteines are modified in eukaryotic sulfatases even after a co-expression with SUMF1 (Cosma et al., 2003; Dierks et al., 2003). Another FGly generating enzyme was identified in prokaryotes, called AstB. This enzyme is responsible for the serine/FGly conversion.

Structures of arylsulfatases were solved by X-ray crystallography from human and from the gram-negative bacteria *Pseudomonas aeruginosa*. They all show similar structure and organization. They have a large N-terminal region comprising 10 mostly parallel β-strands surrounded by α-helices. The C-terminal region is smaller and contains 4 antiparallel β-strands followed by a terminal long α-helix. The conserved active site of the enzyme is present in the center of the protein in a narrow cleft where the FGly and a metal ion are located (Stressler et al., 2016).

No sulfatase has been crystallized in the presence of its physiological substrate yet. Consequently, amino-acids residues implicated in the substrate specificity remain unknown. It is hypothesized that the residues are not located in the narrow cleft, but outside the conserved region (Hanson et al., 2004).

5.1.1. Prokaryotic sulfatases

Prokaryotic sulfatases are present in a soluble form in the cytoplasm or the periplasm. The role of sulfatases in bacteria has been related to sulfate scavenging or to bacteria/host relationships in the context of the human microbiota. Recently, there has been increasing interest on GAG bacterial sulfatases. Three bacterial sulfatases targeting heparin have been characterized in the gram negative *Flavobacterium heparinum* (now called *Pedobacter heparinus*): the 2-*O*-sulfatase (Raman et al., 2003), the 6-*O*-sulfatase (Myette et al., 2009a) and the *N*-sulfamidase that remove the *N*-sulfate of glucosamine (Myette et al., 2009b). This latter study showed the absence within *N*-sulfamidase of key histidines that has been reported as critical to the function of *O*-sulfatases. This suggests differences in the mechanism by which *N*-sulfamidase cleaves nitrogen-sulfur bonds compared to that of *O*-desulfation.

Other recent studies analyzed sulfatases of bacteria living inside the human gastrointestinal tract that relies on host glycan foraging to persist in its host. A first study showed the presence in *Bacteroides thetaiotaomicron* of 3 exosulfatases (Ulmer et al., 2014). The first one is specific of HS 6-*O*-sulfated GlcNAc and shares 57.5% homology with the heparin/HS 6-*O*-sulfatase from *P. heparinus*. Phylogenetic analysis revealed that the glucosamine 6-*O*-sulfatase from *B. thetaiotaomicron* and *P. heparinus* defined a gene cluster containing 61 sulfatase genes. Interestingly, this cluster is composed mostly of genes originating from 38 major gut Bacteroides species (Ulmer et al., 2014).

A second enzyme acts on 6-*O*-sulfated *N*-acetyl galactosamine of CS/DS. This activity was reported previously in *Proteus vulgaris*. However the two enzymes differ on their substrate

specificity, as *B. thetaiotaomicron* 6-*O*-sulfatase works only on the non-reducing end of CS oligosaccharides, whereas *P. vulgaris* 6-*O*-sulfatase acts on reducing end of hexasaccharides (Ulmer et al., 2014). Interestingly, this study also identified the first bacterial endosulfatase enzyme that was active at the polymer level. This enzyme removes sulfate groups in the 4-*O*-position from CS/DS disaccharides to large polymeric chains. This endosulfatase ability makes the enzyme unique and different from the CS 4-*O*-sulfatase characterized in *P. vulgaris* that is only active on the reducing end of oligosaccharides up to hexasaccharides (Ulmer et al., 2014). Indeed, this CS endo-4-*O*-sulfatase shares no significant homologies with the human *N*-acetylgalactosamine-4-*O*-sulfatase ArsB, which is an exosulfatase that removes the sulfate groups present only on the GAG non-reducing end (Ulmer et al., 2014). Compared to eukaryotic endosulfatases, this enzyme lacks the additional HD domain present in Sulfs (see page 63).

The study therefore proposed the implication of this enzyme in the metabolism pathway of host GAGs, suggesting that this step should be the first one during GAG depolymerization (Figure 22, Ulmer et al., 2014). To summarize, CS and DS can be first desulfated by this unique **endo-4-*O*-sulfatase**. Then like HP/HS, they are digested by **lyases** (heparinases for HP/HS and chondroitinases for CS/DS) to generate oligosaccharides with uronic acids at the non-reducing end. These sugars are then processed by the **Δ4,5-hexuronate-2-*O*-sulfatase** and next hydrolyzed into shorter oligosaccharides or monosaccharides by **glycosidases** where each enzyme has its specific substrate. It has been shown for example that the recombinant ∆4,5-glycuronidase preferentially depolymerizes HS/HP rather than CS/DS and/or HA, because it is more efficient on the (1→4) linkage than the (1→3) linkage (Myette et al., 2002). Resulting monosaccharides and oligosaccharides (with a non-reducing end hexosamine) become substrates for the two specific **6-*O*-sulfatases** (galactosamine 6-*O*-sulfatase for CS/DS and glucosamine 6-*O*-sulfatase for HS/HP). Further action of ***N*-sulfamidase** and **3-*O*-sulfatase** are required for HS and HP to achieve complete desulfation. In agreement with this, the action of the 2-*O*-sulfatase in *Flavobacterium heparinum* must precede the Δ4,5-glycuronidase cleavage and the 6-*O*-sulfatase enzyme should act prior to *N*-sulfamidase. However, before the action of these two latter enzymes, that of the 3-*O*-sulfatase must take place (Myette et al., 2009b).

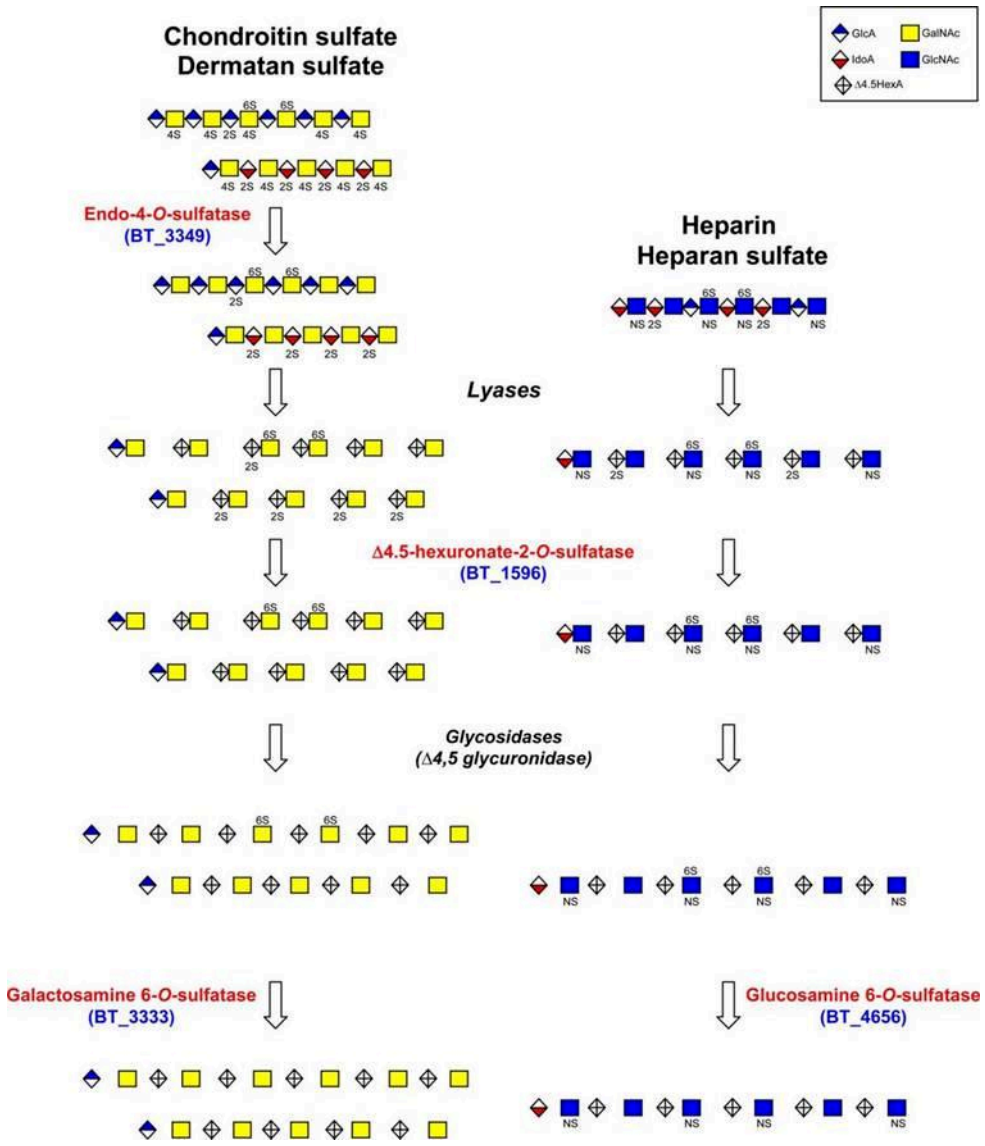


Figure 22: Proposed functions of *B. thetaiotaomicron* sulfatases in the bacterial degradation pathways of glycosaminoglycans. Adapted from (Ulmer et al., 2014).

5.1.2. Eukaryotic sulfatases

Contrary to prokaryotic sulfatases, eukaryotic sulfatases undergo further PTMs. They are first glycosylated, then secreted to cellular compartments or to ECM after cleavage of the signal peptide. The cysteine oxidation takes place in a late translational phase after the translocation of the sulfatase to the ER and before protein folding (Dierks et al., 1997, 1998b; Fey et al., 2001). The final destinations of the sulfatases are either in the Golgi and the ER where they are membrane bound, or in the lysosome and the ECM where they are soluble. The location of sulfatases is correlated with their biological role. For example, lysosomal sulfatases are important for degradation of GAGs and glycolipids, while ER and Golgi sulfatases play a role in the synthesis of hormones. Regarding extracellular sulfatases, they

are important for the control of cell signaling processes. The lysosomal sulfatases act at acidic pH. Among them, ARSA exhibits the ability to desulfate substrates like sulfatides, especially sphingolipids (Mehl and Jatzkewitz, 1968; Roy, 1975). Defect in ARSA results in the accumulation of sulfatide and is at the origin of the metochromatic leukodystrophy (MLD) genetic disorder, which affects the production of myelin in the nervous cells. Other lysosomal sulfatases such as ARSB, galactosamine-6-sulfatase, glucosamine-3-sulfatase, glucosamine-6-sulfatase (G6S), glucouronate-2-sulfatase, heparan-N-sulfatase, and iduronate-2-sulfatase have important roles in the degradation of GAGs. ARSB hydrolyzes sulfate esters at the 4-position of GalNAc residues found in DS and CS (Matalon et al., 1974). Galactosamine-6-sulfatase catabolizes sulfate esters at the 6-position of GalNAc of DS/CS as well as the 6-*O*-S groups found on galactose residues of KS (Bielicki and Hopwood, 1991). The substrates of heparan-N-sulfatase are the *N*-linked sulfamates of glucosamine residues in HP/HS.

In contrast with lysosomal sulfatases, ER and Golgi bound sulfatases work at near neutral pH. They are composed of ARSC, ARSD, ARSF, ARSG for ER sulfatases and of ARSE for Golgi sulfatase. None of them acts on GAGs. ARSC for example targets substrates like idothyronine sulfate. It is important to note that the ARS nomination is due to the ARylSulfatase activity of the enzymes, which is their ability to desulfate aryl compounds such as the commonly used 4-methylumbelliferyl sulfate (4MUS) arylsulfatase pseudosubstrate. Finally, sulfatases found in the extracellular compartment correspond to the recently discovered family of enzymes called Sulfs. Contrary to the other exosulfatases, Sulfs are the only eukaryotic enzymes that exhibit endosulfatase activity and also the only ones to display an additional hydrophilic basic domain termed HD. All sulfatases are thus 500-600 amino-acid proteins, except for the Sulfs, which sequence contains over 800 amino-acids residues (Figure 23).

Human Sulfatases	1*	2	3	4	5	6	7	8	9	10	11	12	Position of FGly (1*)	Length
ARSA	C	T	P	S	R	A	A	L	L	T	G	R	**69**	507
ARSB	C	T	P	S	R	S	Q	L	L	T	G	R	**75**	533
ARSC	C	T	P	S	R	A	A	F	M	T	G	R	**83**	583
ARSD	C	T	P	S	R	A	A	F	L	T	G	R	**89**	593
ARSE	C	T	P	S	R	A	A	F	L	T	G	R	**86**	589
ARSF	C	S	P	S	R	S	A	F	L	T	G	R	**79**	591
ARSG	C	S	P	S	R	A	S	L	L	T	G	R	**82**	525
GalN6S	C	S	P	S	R	A	A	L	L	T	G	R	**79**	522
GlcN6S	C	C	P	S	R	A	S	I	L	T	G	K	**91**	552
GlcNS	C	S	P	S	R	A	S	L	L	T	G	L	**70**	502
IdoA2S	C	A	P	S	R	V	S	F	L	T	G	R	**84**	550
Sulf1	C	C	P	S	R	S	S	M	L	T	G	K	**87**	871
Sulf2	C	C	P	S	R	S	S	I	L	T	G	K	**88**	870

Figure 23: Signature sequences of sulfatases. Partial alignment of sulfatases from all sulfatase genes cloned shows homology of the sulfatase signature sequences. This consensus sequence is important for directing the first amino-acid residue to the catalytically active FGly for oxidation. Highly conserved residues are shown in white letters on a black background; other significantly conserved residues are shown in gray. From (Hanson et al., 2004).

5.2. Sulfs: the state of the art

5.2.1. Discovery of Sulfs

Sulf-1 was first discovered in quail by dhoot and colleagues in a study screening SHH responding activated genes during the development of quail embryos (Dhoot et al., 2001). It was classified as a sulfatase, given that its N-terminal region was homologous to that of lysosomal G6S sulfatase. The importance of this discovery was underlined by the ability of the enzyme to modulate positively the Wnt signaling pathway in muscles progenitors. Since, Orthologs have been identified in mouse, rat, chick, *C. elegans*, zebrafish and in human. In human, Sulfs exist as two isoforms: HSulf-1 and HSulf-2, encoded by two distinct genes, which feature a common structural organization (Morimoto-Tomita et al., 2002). Sulfs are secreted in the extracellular medium and exhibit arylsulfatase activity on 4MUS at neutral pH, as well as endosulfatase activity on HS/HP chains (Morimoto-Tomita et al., 2002). The optimal activity of QSulf-1 is at pH 7.5 and in the presence of Mg_2^+. The enzymatic activity could be upregulated by Pb_2^+ and inhibited by 25 mM phosphate or sulfate (Ai et al., 2003). It is not clear whether these ions are important for human Sulfs as well. Sulfs therefore distinguishes from other sulfatases by being **extracellular** and **endosulfatases** enzymes. They also contain **unique structural organization** due to the insertion of a unique hydrophilic domain HD in their C-terminal region.

5.2.2. Organisation in domains

Sulfs are initially synthetized as pre-pro-proteins, containing a **cleavable signal sequence** of 22 and 24 amino-acids residues for HSulf-1 and HSulf-2, respectively. Once removed in the ER, the 125 kDa pro-proteins are then processed by furin-type proteases, known to be active in several cellular compartments like trans-Golgi network, cell surface and endosome. This cleavage yields mature proteins composed of two subunits linked by one or more disulfide bonds: a 75 kDa N-terminal and the 50 kDa C-terminal regions (Figure 24, Frese et al., 2009; Tang and Rosen, 2009). It is important to note that a band of 240 kDa of Sulf could be detected by Western blot in the extracellular medium, suggesting the existence of oligomeric forms of Sulfs (Tang and Rosen, 2009).

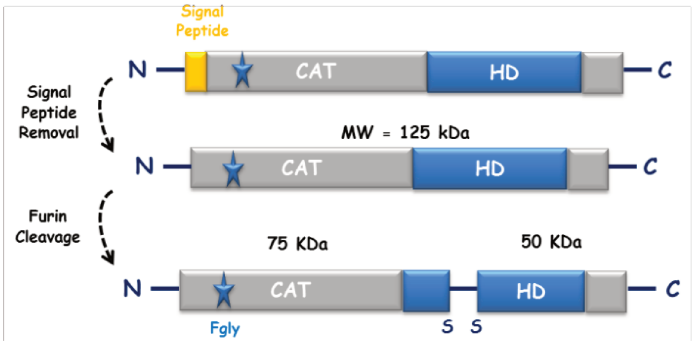


Figure 24: Schematic representation of Sulfs' synthesis and domains.

In the 75 kDa N-terminal region, the catalytic domain (**CAT**), composed of 392-391 amino-acids residues for HSulf-1 and HSulf-2 respectively, shares significant sequence homology with all members of the Sulfs family (81% of homology between HSulf-1 and HSulf-2 CAT domains, see page 176). It comprises the active site of the enzyme, including the two conserved sulfatase signature sequences, and the strictly conserved FGly at position 89 and 88 for HSulf-1 and HSulf-2 respectively, that is essential for arylsulfatase activity. In contrast, next to the CAT is the hydrophilic, highly charged basic domain (**HD**) that is a unique feature of the Sulfs, sharing no homology with any other known protein. It comprises 327 amino-acids for HSulf-1 including of 23% basic residues and 14% of acidic residues, and 307 amino-acids residues for HSulf-2 including 21% of basic and 13% of acidic residues. HD thus confers to Sulfs a theoretical pI around 9.2-9.3, which is thus highly positively charged at neutral pH. Although not necessary for the arylsulfatase activity, the HD domain is responsible for the recognition and high affinity binding to HS substrates and is required for the enzyme endosulfatase activity (i.e. Sulf-specific ability to catalyze HS 6-*O*-desulfation)

(Ai et al., 2006; Frese et al., 2009; Morimoto-Tomita et al., 2002; Tang and Rosen, 2009). Interestingly, HD domains from HSulf-1 and HSulf-2 show poor sequence homology (43%) suggesting the existence of isoform-specific substrate preferences. The conserved regions are located in their outer regions, and especially comprise a cluster of basic residues at the C terminal end of the HD domain, whereas the inner region of the HD is less conserved (see page 176). Finally, the end of Sulf **C-terminal** region, composed of 130-148 residues for HSulf-1 and -2 respectively, is homologous to Glucosamine-6-sulfatase (G6S) and also with GlcNAc transferase from *Arabidopsis thaliana*, suggesting a role of this domain in the recognition of glucosamine motifs (Morimoto-Tomita et al., 2002). It is highly conserved among isoforms (71%, see page 176).

Sulfs are secreted proteins that exert their activity in the extracellular compartment, on both cell-surface and ECM HS. The attachment of Sulfs on cell surface may be due to the interaction between HD domain and cell surface HS. In fact, the deletion of HD from QSulf-1 and HSulf-1 results in the release of the enzymes in the extracellular medium. In addition, it has been shown by FACS, that heparinase treatment of fibroblast prevents the HD of HSulf-1 to bind to the cell surface (Frese et al., 2009). However, heparinase treatment of cells expressing QSulf-1 did not release the enzyme, suggesting thus that binding of the enzyme to the cell surface does not exclusively rely on HS and may involve interactions of the HD with other cell surface components, possibly other types of GAGs (Dhoot et al., 2001). Indeed, although Sulfs bind preferentially to their substrates (HP or HS), HSulf-1 has been shown to weakly bind to non-substrate GAGs like CS/DS and Sulfs pre-treated HS, mainly through the CAT domain (Figure 25, Milz et al., 2013).

The outer regions of HD, which are conserved between isoforms and amongst species, seem to be responsible for the attachment to cell surface HS. The basic cluster in the HD C-terminus alone is not sufficient to mediate attachment (Frese et al., 2009). However, in QSulf, this cluster seems to be not necessary for anchoring to the cell surface (Ai et al., 2006). Deletion of the Sulf-1 HD inner region did neither affect the binding to the cell surface in quail (Ai et al., 2006), nor enzyme activity in human (Frese et al., 2009).

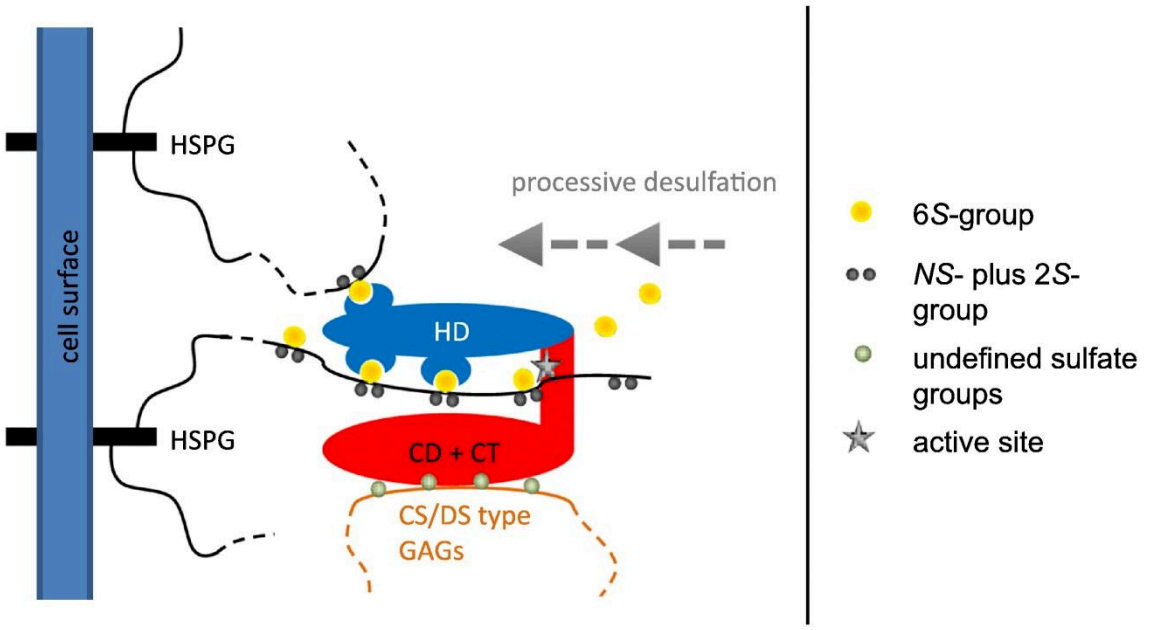


Figure 25: Hypothetical model for a processive cooperation of Sulf-1/GAG binding sites. HD (dark blue) forms an exosite extending from the catalytic (CD) and C-terminal (CT) domains (red) of Sulf-1. HD interacts with heparan sulfate chains showing high specificity for the 6-*O*-sulfate substrate groups (yellow balls) and presents these groups to the active site (gray star) of CD, where they are sequentially removed from the chain. In addition, CS or DS can bind the CD/CT domains *via* undefined sulfate groups (green balls), but are not enzymatically desulfated. From (Milz et al., 2013).

5.2.3. Post-translational modifications

To become mature proteins, Sulfs undergo **PTMs**. The **FGly** conversion from cysteine is the common modification to all sulfatases and is essential to the arylsulfatase activity of Sulfs. In absence of any detailed study, Sulf desulfation mechanism *per se* has been assumed to occur following the general arylsulfatase course (see page 56, Hanson et al., 2004).

As mentioned before, Sulf pro-protein becomes mature after **furin cleavage**. Furin are cellular endoproteases that catalyze the maturation of many secreted proteins implicated in a variety of physiological processes. Sulfs features two consensus sites for furin cleavage located in the inner region of Sulfs HD, which are conserved among isoforms (Nagamine et al., 2010). For HSulf-2, the first one is **R**$_{511}$-S-I-**R**$_{514}$ (with cleavage after the last arginine, see page 174) and corresponds to the most common furin target sequence, while the second one, **R**$_{536}$-N-L-T-**K**-**R**$_{541}$, is a less frequent site. Studies investigating the mechanism and the importance of furin processing led to conflicting data. First, it has been suggested that furin acts on both sites, explaining the presence, in Western blot, of two ~50 kDa fragments corresponding to Sulf 50 kDa C-terminal region (Tang and Rosen, 2009). Noteworthy, alteration of the first site partially blocked the enzyme cleavage, whereas alteration of the

second one did not seem to affect the cleavage. However, furin cleavage was completely blocked when both sites were mutated (Nagamine et al., 2010). In another study, the whole length protein form was detected in the extracellular medium, suggesting that furin cleavage is not required for the enzyme secretion (Tang and Rosen, 2009). In addition, it has been shown that the substitution of the furin cleavage sites did not affect HSulf-2 activity, secretion or solubility, but did affect its location in the lipids rafts of the cell surface, thus decreasing Wnt signaling (Tang and Rosen, 2009). However, it has also been shown that the removal of the furin cleavage site did not affect FGF2 signal transduction (Frese et al., 2009). Noteworthy, the absence of the furin cleavage could increase the occurrence of unprocessed whole length enzyme prone to dimerize. This may be due to the C-terminal domain that tends to dimerize unlike the N-terminal. Further studies will be needed to clarify the role of furin in Sulf maturation.

Sulfs are ***N*-glycosylated** proteins, with 10 or 11 potential *N*-linked glycosylation sites, located mostly on the N-terminal domain and accounting for ~20 % of the protein molecular weight (MW). Although not investigated in human forms yet, a study on Quail Sulf-1 indicated that these glycosylations were necessary for appropriate cell surface localization and enzyme activity (Ambasta et al., 2007).

5.2.4. Desulfation process of Sulfs

Contrary to other sulfatases, Sulfs are endo-enzymes, which preferentially target internal HS highly sulfated NS domains. Sulfs catalyze the 6-*O*-desulfation of [UA(2S), GlcNS(6S)] trisulfated disaccharides units essentially, although residual activity on [UA, GlcNS(6S)] disulfated disaccharides has also been reported (Ai et al., 2003; Pempe et al., 2012; Seffouh et al., 2013; Staples et al., 2011). The absence of activity on GlcNAc containing disulfated disaccharides suggests a requirement for *N*-sulfate groups. In addition, HSulfs seem to indistinctly accommodate both IdoA- of GlcA- containing disaccharides (Pempe et al., 2012), while QSulf shows activity on [GlcA, GlcNS(6S)] but not on [IdoA, GlcNS(6S)] disaccharides (Ai et al., 2003; Viviano et al., 2004).

The process of HS 6-*O*-desulfation is not fully understood yet. Recent data have provided further insights into the underlying mechanisms. HSulfs first bind with high affinity to the polysaccharide through their HD domains. Surface plasmon resonance analysis of HSulf-1 HD domain yielded high affinity Kds (in the nanomolar range) and showed the formation of very stable enzyme-substrate complexes. Interestingly, binding data could not be fitted to a

simple 1:1 binding model, thereby suggesting the existence of a complex mode of interaction (personal data). Importantly, high affinity HS/HD domain interaction requires 6-*O*-sulfation, and occurs through both inner and C-terminal regions of HD domain, suggesting the presence of multiple HS binding sites. The N-terminal part of HD does not seem to contribute to the high affinity binding (Frese et al., 2009; Milz et al., 2013). In a recent study, the unique dynamic properties of HD-HSulf1/HS interactions were analyzed further using atomic force microscopy at the single molecule level. This biophysical approach consists of immobilizing the HD and its counter ligand HS on opposing surfaces and quantitatively investigating their binding properties with single molecule force spectroscopy. The HD-HSulf1/HS interaction appeared to be of catch-bond type under a range force of 10 to 18 pN, as it exhibited increased dissociation lifetime when subjected to external forces (Harder et al., 2015). Outside these ranges, slip bond dissociation was observed, the transition from one type to another being probably due to the system reaching its maximum stability. The catch behavior was associated to the 6-*O*-S groups, given that the interaction between similar substrate lacking these groups (HS precursor *N*-sulfated heparosan K5-NS) displayed a slip type behavior under the full range of force. In addition, it appeared that these interactions showed more than one binding state.

Recently, our laboratory showed that Sulfs-catalyzed desulfation always initiated at the non-reducing end of HS NS domains and proceeded towards the domain reducing end in a processive manner (Seffouh et al., 2013). This implied that Sulf HD domain would primarily bind a saccharide motif downstream the NS domain non-reducing end to fit adequately the first 6-*O*-sulfated glucosamine in the active site. Hydrolysis of the first 6-*O*-S groups released sufficient energy to drive the desulfation process. The free energy landscape of the two-state model was investigated and suggested the existence of additional intermediate states, which would be too transient to be observed, or could only be observed in the full length protein (Walhorn et al., 2018).

We, and others, thereby proposed a model where 6-*O*-S act as allosteric effectors that induce the force mediating a conformational switch transition state, to feed the upstream 6-*O*-S to the catalytic sites and pull the enzyme along the polysaccharide without detaching the GAG chain (Harder et al., 2015; Seffouh et al., 2013). Once the NS domain reducing end is reached and in the absence of downstream sulfated residue, the affinity of the HD domain for the HS chain would drop and the enzyme would be released from the desulfated polysaccharide (Figure 26).

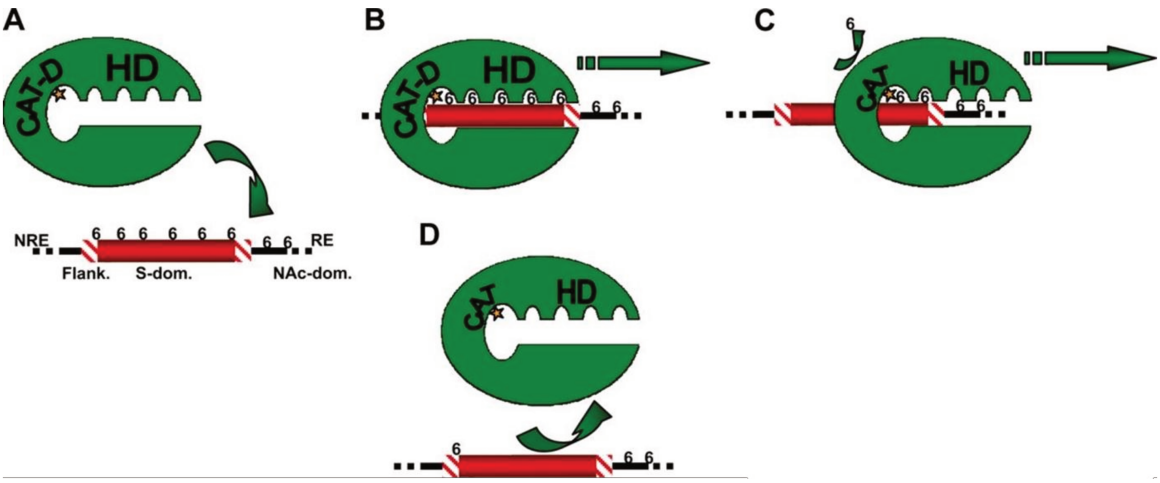


Figure 26: Model for HSulf processive activity and differential regulation of HS ligand binding properties. HSulf HD preferentially recognizes and binds to 6-*O*-sulfates harbored by trisulfated disaccharides found within HS NS domains (A), the catalytic site (CAT-D) of the enzymes being positioned on the most upstream 6-*O*-sulfate residue (B). After desulfation, the enzyme progresses along the polysaccharide chain to accommodate other 6-*O*-sulfate groups (C). Once the NS domain reducing end is reached, the absence of appropriate 6-*O*-sulfates (those present on the flanking regions (Flank.) or the NAc domains (NAc-dom.) are poor substrates for the enzyme) downstream CAT-D results in a strong decrease in the affinity of the enzyme HD for the polysaccharide and the dissociation of the complex. From (Seffouh et al., 2013).

5.2.5. Substrate specificity of HSulf-1 and HSulf-2

Sulfs are still poorly characterized enzymes and major issues remain to be clarified. Many data underlined the existence of possible functional or structural differences between Sulf isoforms that confer the substrate specificity to the enzymes. Among them are the differences in their expression pattern, their *in vitro* activities, and their *in vivo* implications.

The expression of Sulfs is not ubiquitous and varies among tissues in an isoform specific manner. The highest levels of Sulfs expression in adult human were found in testes, stomach, skeletal muscle, lung, and kidney for Sulf-1 and in ovary, skeletal muscle, stomach, brain, uterus, heart, kidney, and placenta for Sulf-2 (Morimoto-Tomita et al., 2002).

In vitro, analysis of HSulf-1 and HSulf-2 activity showed that they were able to catalyze full 6-*O*-desulfation of a heparin oligosaccharide, but with different rates (Seffouh et al., 2013). However, regarding the processive mechanism shared by HSulf-1 and HSulf-2, it may point out the HD domain as the key regulator of the enzyme activity and of isoform preferences.

In vivo, studies on Sulf KO mice also provided indications about the existence of isoform specificities. Sulfs exhibited either redundant, overlapping or even opposite functions,

depending on the biological systems considered. Studies on the role of Sulfs during development and in cancer provided strong evidence of such discrepancies. In mice, the single KO of Sulf-1 or Sulf-2 did not cause severe abnormalities or histological defects. However, mSulf-2 KO mice displayed severe brain malformations and died within 6 weeks after birth. Sulf-1/Sulf-2 double KO animals led to high neonatal mortality and multiple phenotype anomalies along with significant decrease in body mass (Holst et al., 2007; Lum et al., 2007). These data therefore pointed out major overlapping functions and/or compensation effects between the two isoforms during development. In addition, both enzymes played overlapping activity in the change of cell fate from motor neurons to oligodendrocyte precursor cells, by regulating SHH signaling in the ventral spinal cord of mice. However, in this case, no compensatory effect by one enzyme for the loss of the other was observed (Jiang et al., 2017). In agreement with this, analysis of HS from KO mice revealed that 6-*O*-sulfation content was significantly higher in Sulf double KO than in simple KO mice, thereby supporting some functional cooperativity of Sulf-1 and Sulf-2 isoforms (Lamanna et al., 2006). Surprisingly, significant increase of 6-*O*-sulfation was detected in HS NAc/NS transition zones. This was therefore in contradiction with *in vitro* analysis, which reported that Sulfs exclusively targeted highly sulfated disaccharides that are normally present within NS domains. Interestingly, a significant reduction in HS *N*- and 2-*O*-sulfatation content, along with changes in 2OST and 6OST expression were also observed (Lamanna et al., 2008). These data thus suggested the existence of interconnexions between Sulf activity and HS biosynthesis machinery *in vivo*.

Interestingly, increased expression of Sulf-1 could compensate the loss of Sulf-2, but Sulf-2 could not completely substitute for the lack of Sulf-1 (Lamanna et al., 2006). Furthermore, comparative analysis of HS from single Sulf-1 and Sulf-2 KO mouse organs revealed sulfation differences (Nagamine et al., 2012). Non-redundancy of Sulf functions was also reported during mouse brain development, as the two isoforms differently contributed to neurite outgrowth of cerebellar and hippocampal neurons, synaptic plasticity, and motor activity (Kalus et al., 2009). Major differences between Sulf-1 and Sulf-2 activities have also been reported in cancer (see page 72). Although still debated, HSulf-1 has been frequently associated with anti-oncogenic activities, while HSulf-2 has been generally associated with pro-oncogenic activities (Rosen and Lemjabbar-Alaoui, 2010; Vivès et al., 2014). Surprisingly, both HSulfs have been shown to downregulate pro-angiogenic growth factors *in vitro*. However, HSulf-2 promoted tumor angiogenesis *in vivo* (Lai et al., 2008; Morimoto-Tomita et al., 2005). The rationale behind such a discrepancy still remains poorly understood.

However, one explanation could be that the two isoforms feature subtle substrate specificities and/or act on different HS subsets. For instance, targeting cell surface HS that acts as pro-angiogenic/pro-oncogenic growth factor coreceptors, or ECM HS involved in sequestration and storage of these growth factors, may have opposite effects (Figure 27)

5.3. Sulf regulation of HS binding proteins

Because of their stringent substrate preference for [IdoA(2S), GlcNS(6S)] units (which usually account for less than 10 % of HS disaccharide content), consequences of Sulf activity on HS are structurally very limited. However, by specifically targeting HS functional NS domains, Sulfs have been shown to dramatically alter the polysaccharide binding properties, with very variable functional consequences (Figure 27).

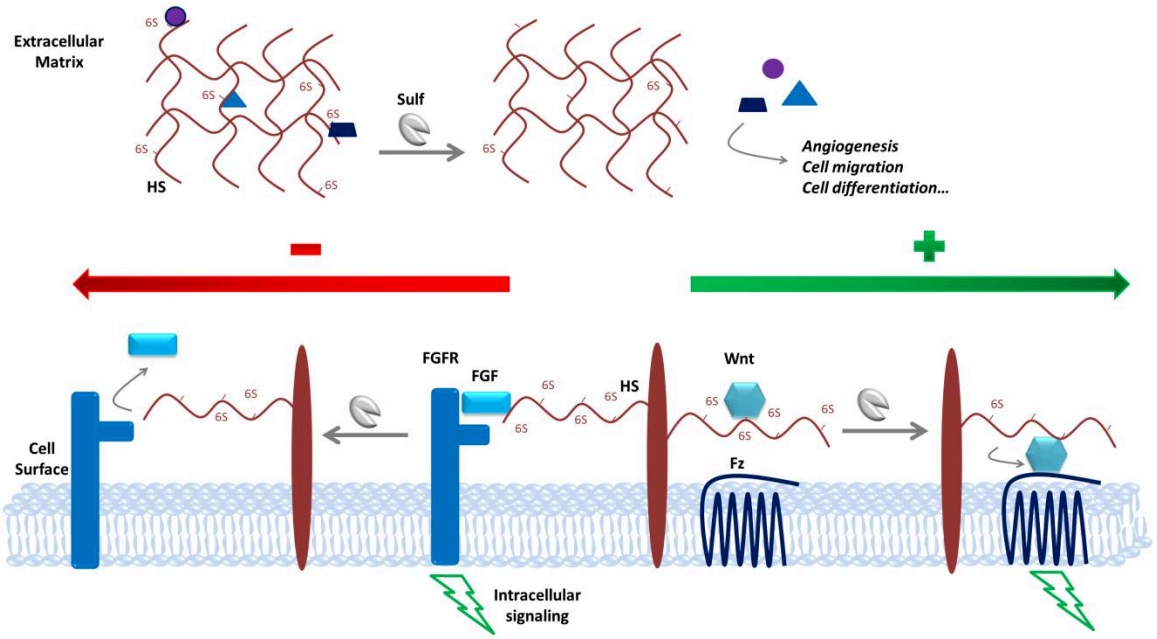


Figure 27: Regulation of heparin binding proteins by Sulfs. Regulation of HP binding protein signaling by the Sulfs. At the cell surface, 6-*O*-desulfation mediated by the Sulfs can have different consequences on the signaling of HS binding proteins. It can downregulate FGFs (left side), by preventing formation of the FGF/HS/FGFR ternary complex that triggers intracellular signaling, and it can induce Wnt (right side), by lowering the affinity of the Wnt/HS interaction, thereby allowing subsequent binding to the Fz receptor. In contrast, alteration of ECM HS structure by the Sulfs may release sequestered proteins (including FGFs) to elicit their functions. From (El Masri et al., 2017).

Sulfs indeed inhibit a variety of HP binding proteins, including growth factors such as FGF1 (Seffouh et al., 2013; Uchimura et al., 2006), FGF2 (Lai et al., 2003; Li et al., 2005; Narita et al., 2006; Wang et al., 2004), HGF (Lai et al., 2004a; Narita et al., 2006), HB-EGF (Dai et al., 2005; Lai et al., 2003), amphiregulin (Narita et al., 2007) or TGFβ (Yue et al., 2008). In many cases, downregulation is the consequence of a structural alteration of cell-

surface HS acting as coreceptors for these growth factors. In addition, compromised binding may also have direct consequences for proteins such as chemokines, which activity relies on the formation of protein gradients stabilized through interactions with PGs. Inhibition of HS interaction with CXCL12 and CCL21 by HSulf-2 has been demonstrated *in vitro* (Uchimura et al., 2006). Although the physiological relevance of such effect has not been investigated yet, effects on leukocyte recruitment/homing may be anticipated.

In contrast, Sulfs have also been shown to induce signaling pathways, as originally demonstrated for Wnt morphogen (Dhoot et al., 2001). As explained above, Sulfs mediated 6-*O*-desulfation weakens Wnt/HS interaction, allowing formation of a ternary HS/Wnt/Fz signaling complex (Ai et al., 2003). A similar mechanism has also been reported for the regulation of GDNF (Ai et al., 2007; Langsdorf et al., 2011). Finally, Sulfs have been shown to indirectly promote BMP signaling, by modulating its inhibition by Noggin (Otsuki et al., 2010; Viviano et al., 2004). Unexpectedly, Sulfs have demonstrated opposite activities for some ligands. For instance, Sulf-1 has been shown to either inhibit or induce SHH signaling in gastric cancer or during neuronal development, respectively (Danesin et al., 2006; MA et al., 2011). In addition, Sulf-1 appears to induce TGFβ1 in mice intervertebral disc homeostasis (Otsuki et al., 2019) , and to inhibit its expression in the development of pulmonary fibrosis (Yue et al., 2008).

Thanks to these complex HS regulatory properties, Sulfs are involved in major physiological processes. Notably, Sulfs have been shown to play central roles during development, which could be partly deduced from the study of KO animals. In mouse and chick, Sulfs have been implicated in neuronal (Ai et al., 2007; Danesin et al., 2006; Kalus et al., 2009, 2015; Oustah et al., 2014), skeletal and cartilage (Holst et al., 2007; Zhao et al., 2006) development, formation of the inner ear (Freeman et al., 2015) and dentinogenesis (Hayano et al., 2012). In *C. elegans*, Sulf-1 has been shown to participate to the dorsal ventral patterning of the neural tube (Ramsbottom et al., 2014; Winterbottom and Pownall, 2009). Finally, in zebrafish, alteration of BMP, FGF and CXCL12 signaling resulting from the loss of Sulf-1 led to poor differentiation of the somitic trunk muscle, loss of the horizontal myoseptum, reduction of pigmentation along the mediolateral stripe, and incorrect migration of the lateral line primordium (Meyers et al., 2013). In adults, Sulfs have been implicated in tissue and organ regeneration and wound repair (Maltseva et al., 2013; Nakamura et al., 2013; Tran et al., 2012; Yue, 2017), cartilage homeostasis and intervertebral disc homeostasis (Otsuki et al., 2010, 2019), changing cell fate from motor neurons to oligodendrocytes

precursor cells (Jiang et al., 2017) as well as maintenance and stability of renal glomerular filtration barrier (Schumacher et al., 2011; Takashima et al., 2016). Sulfs have also been very early associated with a number of diseases.

5.4. Sulfs in Cancer

Sulfs have been associated with large variety of cancers (leukemia, ovarian, liver, pancreas, lung, breast, brain, kidney, bladder, colon, gastric, and head and neck cancer) and have been involved in all major stages of the disease, including tumoral transformation (Rosen and Lemjabbar-Alaoui, 2010), growth, invasion and metastatization (Abiatari et al., 2006; Khurana et al., 2012a, 2013a; Li et al., 2005; Nawroth et al., 2007; Phillips et al., 2012), as well as tumor cell sensitization/resistance to drugs (Lai et al., 2003; Moussay et al., 2010). HSulf-2 in particular has been marked as a target of interest in cancer therapy, especially for tumors of poor prognosis, such as lung squamous cell carcinoma and lung adenocarcinoma (Lemjabbar-Alaoui et al., 2010; Rosen and Lemjabbar-Alaoui, 2010). Interestingly, HSulf-1 and HSulf-2 appear to have divergent activities. HSulf-1 has an anti-oncogenic activity, while HSulf-2 has a pro-oncogenic role in most cancers (Vivès et al., 2014).

5.4.1. Sulf-1

A decrease in Sulf-1 expression has been found in many cancer cell lines like breast, pancreas, kidney, liver and ovarian cells and in numerous cancer specimens such as hepatocellular, breast, gastric, renal and colon cancers, especially in the early stage of cancers (Ji et al., 2011; Lai et al., 2003). .

In **hepatocellular carcinoma** (HCC), which is one of the highest mortality cancers for being diagnosed at late stage, Sulf-1 acts as a tumor suppressor. A downregulation of mRNA HSulf-1 was observed in primary hepatocellular carcinoma (30%) and more pronounced in HCC cell lines, occasionally accompanied with allelic loss. For the 70% samples where HSulf-1 is up regulated, it has been suggested that HSulf-1 was co-amplified with the myc gene, given that HSulf-1 loci is located near that of myc, which is up regulated in HCC cancers (Lai et al., 2004a). Moreover, it has been shown that overexpression of Sulf-1 in HCC murine cancer model led to a decrease in cell proliferation, division, migration and invasion *in vitro*, and to a reduction of tumor progression and lymph node metastasis *in vivo* (Mahmoud et al., 2016). This effect was explained by the downregulation of mesothelin at both mRNA and protein levels in HSulf-1 overexpressing HCC cells. Furthermore, the same authors showed that Sulf-1 knockdown upregulated mesothelin expression and enhanced

tumor growth and ability to metastasize (Mahmoud et al., 2018). Mesothelin is a cell surface glycoprotein reported to activate Mitogen-activated protein kinases (MAPK) proliferation signaling, thus promoting tumor growth (Chang et al., 2009). It has been suggested that Sulf-1, by removing the 6-*O*-S groups of HSPGs such as syndecans, inhibited the formation of HS/Wnt-1 complex, thereby preventing the release and induction of mesothelin and subsequent downstream proliferation signaling (Figure 28).

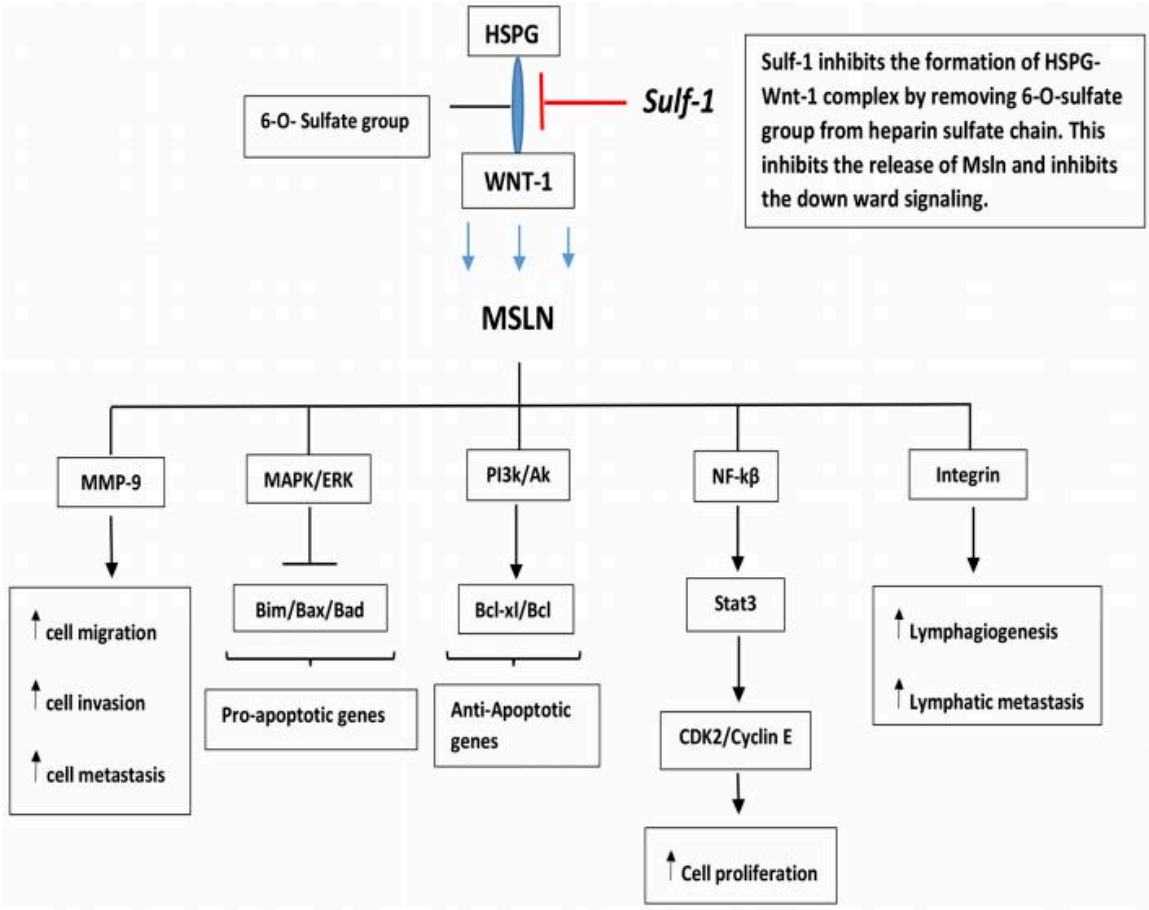


Figure 28: Schematic presentation showing the role of Sulf in cancer. From (Mahmoud et al., 2016).

Furthermore, re-expression of HSulf-1 by adenovirus in HCC cells decreased the phosphorylation of a serine/threonine-protein kinase (AKT) and extracellular signal-regulated kinases (ERK), suppressing cell migration and proliferation (Liu et al., 2013). In addition, expression of HSulf-1 in HSulf-1 negative HCC cells and HCC xenografts reversed the action of exogenous FGFs, through the suppression of AKT and ERK signalings and subsequent downregulation of their targeted genes *cyclin* and *survivin*, resulting in the inhibition of cell cycle progression and induction of apoptosis. From these data, a therapy combining recombinant HSulf-1 and rapamycin (mTOR inhibitor) against HCC cells has been suggested (Xu et al., 2014).

HSulf-1 has also been related to regulation of the MAPK signaling pathway through epigenetic mechanisms. In mice, injection of HCC cells transfected with HSulf-1 resulted in reduced tumor size. It has been shown *in vitro* that HSulf-1 overexpression induced nuclear histone H4 acetylation by changing the balance between histone deacetylase (HDAC) and histone acetyl transferase (HAT) activities, resulting in downregulation of phosphoinositide-3-kinase (PI3) and MAPK kinase pathways, and leading to cell apoptosis and tumor growth inhibition. More interestingly, HSulf-1 promoted the role of HDAC inhibitors (apicidin) during tumor growth and angiogenesis *in vitro* and *in vivo*. HDAC inhibitors are known to activate the apoptosis of cancer cells and are studied in clinical trials for cancer therapy (Lai et al., 2006).

HSulf-1 is also downregulated in primary **ovarian cancer** specimens (as shown by RT-PCR and immunohistology, Lai et al., 2003; Liu et al., 2009). The re-expression of HSulf-1 in ovarian cell types resulted in altered HB-EGF and FGF2 signaling (not EFG) and in increased cell sensitivity to apoptic signaling molecules such as cisplatin. These observations suggested that downregulation of Sulf-1 was a mechanism by which cancer cells could promote tumor growth (Lai et al., 2003). Interestingly, patients with advanced stage ovarian cancer expressing high levels of HSulf-1 responded more efficiently to chemotherapy (Staub et al., 2007). This study suggested a therapy combining epigenetic remodeling and chemodrugs for patients with ovarian cancer exhibiting low level of HSulf-1 expression (Staub et al., 2007). Noteworthy, some polymorphism of Sulf-1 gene appeared to be associated with early stages of ovarian cancer, and could play important role in the prognosis and thereby in the survival of patients (Han et al., 2011).

HSulf-1 has been shown to regulate the glycolysis metabolism in different cancer cells (ovarian, prostate, lung, breast) by modulating the activity of the implicated glycolytic enzymes and the glucose uptake rate. In addition, it modulates the mitochondria function and morphology (Mondal et al., 2015). This implication in mitochondrial metabolism can be explained by the fact that HSulf-1 regulates HB-EGF signaling and c-myc activation, which is a direct regulator of metabolism. In line with this, synthetic agent PG545 has been tested as a substitute of HSulf-1 function in ovarian cancer where HSulf-1 expression is lost (Mondal et al., 2015). PG545 is currently analyzed in phase Ib clinical trials for its ability to mimic HS, sequester growth factors and thereby block angiogenesis. Results showed that PG545 inhibits the ERK, c-myc signaling, and thus the glycolytic enzymes and glucose uptake rate, resulting in the decrease of tumor growth and metastasis. PG545 may therefore act as a new treatment

that reverses glycolytic metabolism alterations in ovarian cancers where HSulf-1 is downregulated (Mondal et al., 2015). HSulf-1 also regulates lipid metabolism in ovarian cancer cells. Loss of HSulf-1 expression in these cells was shown to induce lipid synthesis *in vitro*, thereby facilitating cell proliferation and survival (Roy et al., 2014a). Moreover, the loss of HSulf-1 in ovarian cancer cells promoted anchorage of independent colonies on soft agar *in vitro* and xenograft formation *in vivo*. It has been suggested that enhanced tumorigenesis resulted from the downregulation of pro-apoptic protein Bim caused by the loss of HSulf-1. This decrease in Bim expression was attributed to a degradation of the protein induced by the ERK activating pathway. Interestingly, re-expression of Bim retarded the tumor growth (Figure 29, He et al., 2014).

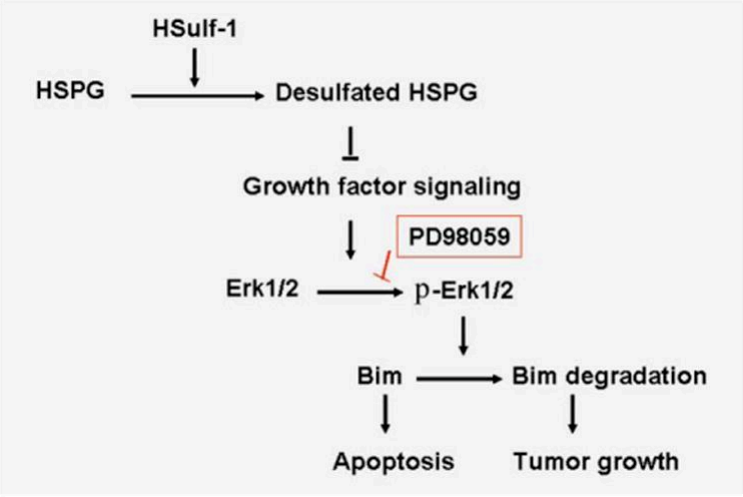


Figure 29: Representation of the effect of HSulf-1 on Bim. From (He et al., 2014).

It has also been shown that the decrease of HSulf-1 was accompanied with an increase of VEGFR phosphorylation in HCC, with no change in expression. Interestingly, re-expression of HSulf-1 in HCC and ovarian cancer cells or xenografts downregulated VEGFR phosphorylation, resulting in inhibition of cell proliferation and ability to metastasize (Ji et al., 2011).

The role of HSulf-1 in reducing the tumor growth was also shown in MDA-MB-468 **breast cancer** cells (Narita et al., 2006). Overexpression of HSulf-1 inhibited angiogenesis and induced cell apoptosis and necrosis (resulting from poor supply in nutrients and oxygen). The overexpression of HSulf-1 in breast cancer cell reduced autocrine EGFR-ERK1/2 signaling mediated by amphiregulin and HB-EGF, without affecting ligands levels. Such altered signaling was associated with the inhibition of cell entry into S phase, resulting in cell

death. Interestingly, the loss of HSulf-1 showed an increase in both autocrine and paracrine proliferation through the same mechanism (exogenous amphiregulin). Levels of mRNA HSulf-1 was reduced in primary invasive breast tumors (60%) associated with hormone-induced estrogen receptor (ER) activation, which is known to promote EGFR (Narita et al., 2007). Furthermore, HSulf-1 regulates Hypoxia-Induced Factor (HIF1α)-mediated FGF2 signaling, migration and invasion of in breast cancer cells (MCF7) *in vitro* (Khurana et al., 2011).

In human **lung epithelial cancer** cell lines, the mRNA level of HSulf-1 is also reduced. Overexpression of HSulf-1 in lung cancer cells decreased proliferation, viability and increased apoptosis, when compared to HSulf-1 overexpression in normal cells. One mechanism responsible for this may involve the reduction of AKT and ERK signaling (Zhang et al., 2012a).

HSulf-1 can also inhibit the **gastric cancer** cell proliferation as well, by downregulating HH signaling (MA et al., 2011). Moreover, re-expression of HSulf-1 in specific gastric cancer cell line reduced the proliferation and invasion, through suppression of the Wnt/β-catenin signaling pathway. The anti-tumor growth effect of HSulf-1 was also confirmed *in vivo* (Li et al., 2011).

HSulf-1 is downregulated in **head and neck squamous carcinoma** (SCCHN) cell lines. Expression of Sulf-1 in SCCHN cell line attenuated HS 6-*O*-sulfation at the cell surface, decreasing FGF2 and HFG signaling and thus resulting in reduced cell proliferation invasion and metastasis. HSulf-1 also increased sensitivity to apoptosis (Lai et al., 2004b).

Finally, endogenous HSulf-1 decreased the ability of **pancreatic cancer** cells to proliferate *in vivo*. Noteworthy, overexpression of HSulf-1 in these cells increased their invasive potential (Abiatari et al., 2006). Consistently, HSulf-1 is upregulated in pancreatic cancer in either early or late stages (Li et al., 2005; Nawroth et al., 2007), as well as in patients with lung tumors (Lemjabbar-Alaoui et al., 2010). Interestingly, these distinctive activities of HSulf-1 were also observed in another study, where a novel Sulf-1 variant (sulf-1B) resulting from gene alternative splicing, which encodes for a shorter protein form, was identified. In fact, contrary to Sulf-1, Sulf-1B inhibited Wnt signaling and promoted angiogenesis (Sahota and Dhoot, 2009).

5.4.2. Sulf-2

The role of Sulf-2 was assessed in many types of cancers. First, the expression of HSulf-2 was analyzed in patients with **renal carcinoma** (RCC). Results showed that the level of HSulf-2 mRNA was significantly higher in cancer tissues than in normal tissues. However, in advanced cancer, mRNA HSulf-2 levels appeared to be higher in normal tissues, but the enzyme could not be detected at the protein level. Increased mRNA may be due to an overreaction of normal tissues against cancer to provoke inflammation or the beginning of metastasis development in normal tissues. Interestingly, patients at advanced clinical stage exhibited low level of HSulf-2 expression, whereas those at non aggressive stage showed high level of HSulf-2. A rationale for this could be that low levels of HSulf-2 expression in RCC cells have been shown to enhance activation of VEGF and FGF signaling (*via* their interaction with HS) *in vitro*, resulting in increased tumor growth and metastasis. Expression of VEGF was found to be higher in cells with low Sulf-2 expression. In addition, HSulf-2 downregulated Wnt signaling in RCC (Kumagai et al., 2016).

Furthermore, the level of Sulf-2 transcripts is increased in mouse **brain glioma** (Johansson et al., 2005). In this study, Sulf-2 expression was found elevated in most tumors, whereas Sulf-1 expression remained unchanged. Sulf-2 protein was detected in 50% of primary human glioblastoma (GBM) tumors. Growth of human GBM cancer cells was prevented by knockdown of Sulf-2, and rescued by overexpressing mSulf-2. In support to this, an *in vivo* study in mice showed that Sulf-2 was highly expressed in invasive gliomas, and caused enhanced tumor growth and shortened survival. This could be due to activation by Sulf-2 of many RTKs, including PDGFR, in invasive glioma (Phillips et al., 2012).

Moreover, Sulf-2 (and Sulf-1) is upregulated in human **pancreatic cancer**. It activates Wnt signaling in pancreatic adenocarcinoma tumor cells *in vitro*. Silencing Sulf-2 inhibited tumor growth *in vivo*. These data suggest that Sulf-2 promotes tumor growth by enhancing Wnt signaling (Nawroth et al., 2007).

HSulf-2 is upregulated in **HCC** and HCC cell lines as well, and this correlates with poor prognosis in HCC patients. The expression of Sulf-2 in HCC cell lines is positively correlated with cell proliferation and migration. Sulf-2 increases binding of FGF2 to HCC cells and promotes its activity. One proposed mechanism is through promotion of glypican-3 HSPG expression at the cell surface. In agreement with this, it has been shown *in vivo* that Sulf-2 promotes tumor growth by inducing glypican-3 expression (Lai et al., 2008).

A study on a cohort of patients with **oesophagal cancer**, using immunohistochemistry, showed that HSulf-2 was found in almost all the specimens, with percentage and staining intensity being higher in squamous cell carcinoma samples than in adenocarcinoma samples. Interestingly, the survival rate correlated with high Sulf-2 expression. Sulf-2 as a secreted enzyme could thus represent an interesting biomarker for the diagnosis and prevention of this poor prognosis type of cancer. In this context, suitable methods for the detection of Sulf-2 in blood or other body fluids would be needed (Lui et al., 2012) .

HSulf-2 as well as HSulf-1 transcripts are induced in patients with **lung tumors** (adenocarcinoma and squamous cell carcinoma) and in non-small cell lung cancer (NSCLC) cell lines. Interestingly, both transcripts were not expressed in the same cell lines. Sulf-2 was produced by 10 out of 10 squamous cell carcinoma samples, but was not found in adenocarcinoma samples. The knockdown of HSulf-2 in lung cancer cells reduced cell growth *in vitro*, tumor progression *in vivo*, and inhibited autocrine Wnt signaling. Surprisingly, the overexpression in non-malignant bronchial epithelial cells showed phenotype changes into transformed cells *in vitro*, but did not form tumors *in vivo* (Lemjabbar-Alaoui et al., 2010). In line with this, methylation of Sulf-2 promoter was associated with survival prognosis in NSCLC patients. This methylation was shown to induce IFN-dependent gene expression. One explanation for this is that the methylation silencing Sulf-2 would increase cell surface HS 6-*O*-sulfation, which may promote to binding of IFN to the polysaccharide, its protection against degradation and consequently the induction of IFN-dependent gene transcription. One of these induced genes is *ISG15*, which increases sensitivity to topoisomerase inhibitors (Tessema et al., 2012).

HSulf-2 expression was also analyzed in patients with **breast cancer** cells. High HSulf-2 expression was mostly observed in metastastic tumors (54%), and in a few cases within primary tumors (8%), associating thus HSulf-2 with tumor progression and metastasis. This high expression was suggested to be the cause of cancer cell survival and migration (Khurana et al., 2013a). Expression levels appeared to be higher in estrogen receptor positive tumors. HSulf-2 was found in a number of breast cancer cell lines, such as MCF-7, BT-20, and BT-549, with a 75 kDa band being detected in the conditioned medium (CM) of MCF7 by Western blotting (H2.3 antibody). Sulf-2 was not detectable in the mammary gland of normal mice but was present in the mammary hyperplastic tissues (2/4 mice) and more significantly in mammary tumors (4/4 mice). In hyperplastic tissues, Sulf-2 was located on luminal epithelial cells, but not on myoepithelial cells. In tumor tissues, it was located in epithelial

cells derived from the tumors (Morimoto-Tomita et al., 2005). The pro-angiogenic properties of Sulfs were demonstrated using the Chick chorioallantoic membrane assay. To do this, 10 day chicken embryo chorioallantoic membranes were treated with different amounts of Sulf-2 (25, 50 or 100ng), and the number of blood vessels brunch points was counted three days later. The 50 ng condition showed the highest activity, as 100ng of VEGF, used as a positive control. Interestingly, when 100ng of Sulf2 were used, the activity was reduced, demonstrating thus a dose dependent effect (Morimoto-Tomita et al., 2005). Moreover, shRNA Knockdown of HSulf-2 in breast cancer cell line MCF10DCIS decreased the tumor size *in vivo*, enhanced apoptosis at the center of the tumor, and resulted in less invasive phenotype, most likely by reducing metalloproteinases (MMP)9 expression and activity, which prevented basement membrane degradation (Khurana et al., 2012a).

In addition, in a breast cancer model in mice, subcutaneous injection of MDA-MB-231 cells transfected with HSulf-2 resulted in the increase of the tumors (Zhu et al., 2016). In contrast, one study showed an anti-oncogenic role of HSulf-2 in the same model. In fact, injection of MDA-MB-231 cells transfected with either HSulf-1 or HSulf-2 resulted in inhibition of tumor growth compared to non-transfected cells. In the early stages, tumors displayed similar volumes, but at the end of the experiment, size of Sulf expressing tumors was found to be significantly smaller compared to untransfected controls. Interestingly, the administration of recombinant purified HSulf-2 in non-transfected tumors did not reduce tumor size. An explanation for this could be that HSulf-2 effects on the tumor or the microenvironment are most significant in the early stages of tumor growth or maybe the effect of HSulf-2 is in a dose dependent manner (Peterson et al., 2010).

Another study also showed the inhibitory role of HSulf-2 in **myeloma cancer**. HSulf-1 and HSulf-2 transfected myeloma cells were injected into SCID mice and showed that both Sulfs suppressed tumor growth and progression. Interestingly, this effect was not seen *in vitro*, indicating that the effect of Sulfs was related to the tumor microenvironment. Indeed, although the authors showed that Sulfs could only remodel HS at the tumor cell surface and not that from the ECM, more HS and more collagen fibrils were found within the Sulfs producing tumors (Dai et al., 2005).

5.5. Sulfs in diseases

Increased expression of both Sulf-1 and Sulf-2 has also been reported in osteoarthritic and aging cartilage (Otsuki et al., 2010). In osteoarthritis, abnormal chondrocyte activation and

cartilage degradation may result from Sulf-catalyzed alteration of HS structure and binding properties, and subsequent effects on the signaling of many HP binding growth factors. More recently, eventhough expression of Sulf-1 is induced in degenerative intervertebral disc cells (Tsai et al., 2015), it has been found to contribute to the intervertable disc development and to maintain its homeostasis (Otsuki et al., 2019). Although a direct link between the enzyme expression and the pathology remains to be demonstrated, this study emphasized further the role of Sulfs in cartilage homeostasis and disease. Another emerging but yet poorly explored area is the implication of the Sulfs during inflammation. As explained before, HS 6-*O*-sulfation is critical for the binding of many chemokines and Sulfs may therefore regulate these interactions, as demonstrated *in vitro* for CXCL12 and CCL21 (Uchimura et al., 2006). HS 6-*O*-sulfation is also required for binding to L-selectin, which is implicated in the early events of leukocyte extravasation (Wang et al., 2002), and has been associated with heparin-induced leukocytosis, through disruption of Selectin- and CXCL12-mediated leukocyte trafficking (Zhang et al., 2012b). This suggests that Sulfs could act as a modulator of leukocyte migration and adhesion to activate the endothelium. In line with this, an *in vivo* study on renal allograft biopsies showed that Sulf-1 expression was repressed in inflammatory conditions (Celie et al., 2007). In another study, Sulf-2 has been implicated in idiopathic pulmonary fibrosis, presumably, through a regulation of TGFβ1 signaling in type 2 alveolar epithelial cells (Yue et al., 2013). TGFβ1 is a major actor during the pathogenesis of pulmonary fibrosis (Yue et al., 2010), which binding to HS requires 6-*O*-sulfation and has been shown to induce both Sulf-1 and Sulf-2 expression in lung fibroblasts (Yue et al., 2008, 2013) and renal epithelial cells (Alhasan et al., 2014). Control of TGFβ1/Sulf expression and activities may thus occur through negative feedback loop system. Further study of such regulatory mechanism could provide novel insights into the role of Sulfs in other TGFβ1 involving processes, such as cell differentiation, chemoattraction, and the control of the balance between cell survival and apoptosis. A more recent study showed that Sulf-2 expression in type II alveolar epithelial cells played an important role in the protection from epithelial lung injury, inflammation and mortality (Yue, 2017).

Finally, Sulf-2 has been associated with Type-2 diabetes mellitus, the enzyme inhibiting very-low-density lipoprotein (VLDL) binding to HS hepatocytes and disrupting triglyceride clearance (Hassing et al., 2012). Interestingly, Sulf-2 was also found overexpressed in the serum of cirrhotic patients, suggesting potential use as serologic biomarker (Singer et al., 2015). More recently, Sulf-1 genetic polymorphism has been associated with *in vitro*

fertilization (IVF) failure (Taghizadeh et al., 2015; Zahraei et al., 2014). In contrast, Sulfs have been suggested to have protective effects in pathological conditions such as Alzheimer diseases and kidney amyloidosis. In these diseases, the accumulation of HS NS domains in kidney and in cerebral amyloids β plaques promotes the formation of toxic amyloid fibrils. The enzymatic remodeling of HS by the Sulfs may therefore prevent this accumulation (Hosono-Fukao et al., 2012; Kameyama et al., 2019).

5.6. Regulation of Sulfs

A modification of the expression of the Sulfs has been found in cancer and in other diseases, but the mechanisms responsible for these regulations remain unclear. Studies suggested several ways. First, epigenetic seems to play important roles in the control of Sulf expression. DNA methylation within exon 1A and histone H3 modifications (deacetylation, methylation) resulted in inactivation HSulf-1 expression in ovarian cancer (cell lines and tumors from patients) (Staub et al., 2007). In support with this, DNA methylase inhibitor treatment of primary hepatocellular carcinoma, in which HSulf-1 mRNA is downregulated, resulted in induction of HSulf-1 expression (Lai et al., 2004a). Moreover, serum samples of patients with breast and gastric cancers showed higher methylation levels of HSulf-1 promoter compared to those from control patient samples. This suggested hypermethylation as a mechanism for the downregulation of HSulf-1 expression in these tumors (Chen et al., 2009). Furthermore, HSulf-2 is silenced by the hypermethylation of its CpG island promoter in lung and gastric adenocarcinoma. This may increase sensitivity to chemotherapy, leading to better patient survival (Wang et al., 2013). As mentioned before, the silencing of Sulf-2 in NSCLC patients may be mediated by methylation of Sulf-2 promoter (Tessema et al., 2012).

Second, cytokines can also control the expression of Sulfs. It was shown that TGFβ1 cytokine can induce the expression of Sulf-1 in lung fibroblasts at the transcriptional level. It has been suggested that this could be controlled by unidentified newly synthetized transcriptional suppressors. Once expressed, Sulf-1 can act as a negative regulator of TGFβ1 activity in pulmonary fibrosis (Yue et al., 2008). Interestingly, the overexpression of Sulf-1 can be accompanied with an increase or decrease of Sulf-2 in a murine model, depending on the cell types. This shows a compensatory effect of the expression among Sulfs isoforms (Yue et al., 2008). In line with this, TGFβ1 induced the expression of Sulf-2 in hyperplastic type II alveolar epithelial cells (AEC), in lung tissues of patients with idiopathic pulmonary fibrosis (IPF). Sulf-2 act then as a positive regulator of TGFβ1 (Yue et al., 2013). It is important to note that TGFβ1 activity depends on the *N*-, 2-*O*- or 6-*O*-sulfate groups of HS

(Lyon et al., 1997). Moreover, in the context of inflammation, TNFα cytokine induces the expression of Sulf-1 in human fibroblasts, resulting in cell proliferation (Sikora et al., 2016).

Third, transcription factors and tumor suppressor genes are also able to regulate Sulf expression. Using gene expression profiling, ChIP assays and transcription factor binding site prediction, it has been shown that tumor suppressor p53 binds and activates the transcription of Sulf-2 (Chau et al., 2009). In case of DNA damage, p53 responds by directly upregulating the transcription of HSulf-2. When Sulf-2 is suppressed, the senescence of cells as a response to stress decreased (Chau et al., 2009).

Moreover, Von Hippel-Lindau (VHL) is a tumor suppressor gene that positively regulates the expression of Sulf-2 by the degrading hypoxia inducible factors (HIF). Once VHL is mutated in renal cancer cell lines, HSulf-2 expression is inhibited at both mRNA and protein levels (Khurana et al., 2012b). Consistent with this, hypoxia can downregulate HSulf-2 expression at mRNA and protein levels, and the knockdown of HIF factors restores HSulf-2 expression, independently of VHL. This regulation occurs through specific sequence, found in HSulf-2 promoter, called hypoxia response element (Khurana et al., 2012b). Indeed, HSulf-1 is also negatively regulated by HIF1α-induced hypoxia in breast cancer cells (MCF) *in vitro* (Khurana et al., 2011).

In addition, it has been reported in ovarian cancer, that the suppression of HSulf-1 mRNA was mediated by variant hepatic nuclear factor 1 (vHNF). It was shown by chromatin immunoprecipitation and luciferase reporter assays that vHNF bound specifically to HSulf-1 promoter on vHNF1-responsive elements located upstream the transcription initiation site, decreasing thus its activity (Liu et al., 2009).

Finally, it has been shown *in vitro* that anoïkis downregulates Sulf-2 expression in breast cancer cell lines (MCF10AT1 and MCF7) which leads to cell death (Khurana et al., 2013a). Anoïkis is a specific form of apoptosis, which is due to a lack of interaction between the cell and the ECM: detachment of cell-surface integrins from ECM by matrix MMPs triggers cell death signal. Matrix detachment is a feature of many cancers, and the acquired resistance of cells to anoïkis apoptosis allows them to migrate towards distant organs.

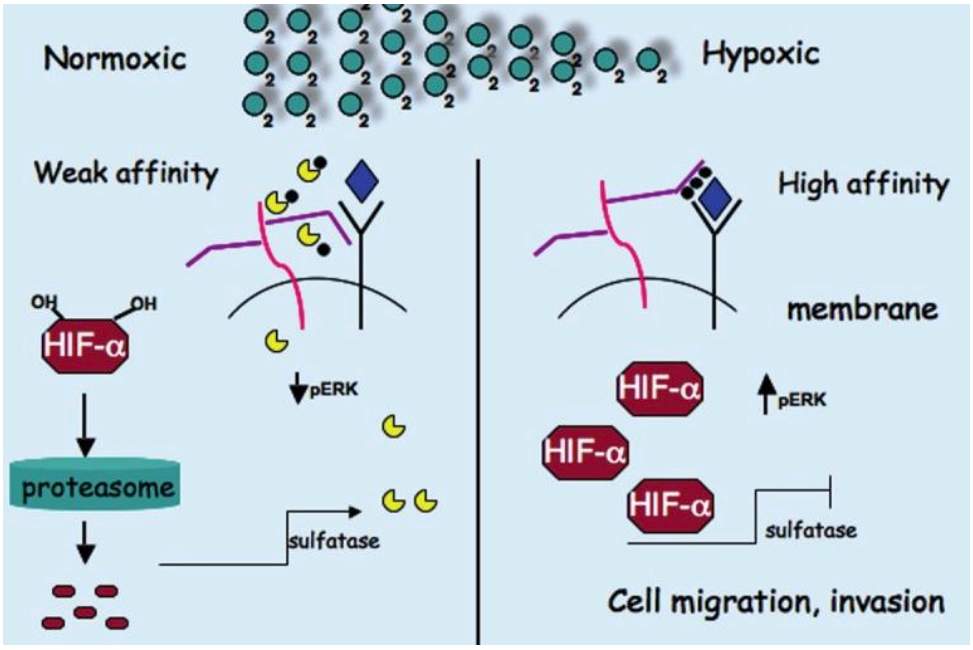


Figure 30: Proposed model of regulation of HSulfs under hypoxic conditions. HSulfs catalytically removes sulfate moiety from 6-*O*-sulfated HS on HSPGs. Desulfation of HS results in decreased FGF2 binding sites on co-receptors (HSPGs) and hence decreased signaling. However under low oxygen conditions (a prevalent condition in solid tumors) or when VHL is inactive, HIF-1α is stabilized and shuts down the transcription of HSulfs and decreased its levels resulting in increased sulfation of 6-*O*-sulfated HS on HSPGs. This increased 6-*O*-sulfation state favors FGF2 signaling, cell migration and invasion. From (Khurana et al., 2013b).

CHAPTER

II

Objectives of the PhD

.

1. The uncharted area of Sulfs

Despite their increasing importance and their implication in many physiological and pathological processes, Sulfs are still poorly characterized enzymes. This is because of their unique properties but also the complexity of their HS substrates. One major problem that has hindered progress in the study of Sulfs is the difficulty to obtain recombinant enzymes. Therefore, little is known about their structure, the importance of their PTMs, their substrate specificities, and their mechanisms. In addition, previously published studies showed confusing data where both isoforms exhibit similar enzyme activity *in vitro* but redundant, overlapping or opposite functions *in vivo*, depending on the biological process considered. These divergent activities were especially emphasized in the field of cancer, where HSulf-1 shows mostly anti-oncogenic properties, whereas Sulf-2 is generally reported as pro-oncogenic.

The objectives of my PhD project were to characterize in details the structural and the functional properties of the human isoform HSulf-2.

Functionally, we have focused our study on two aspects.

1. Clarifying the catalytic mechanism of HSulf enzyme activity *in vitro*. Recently, our group showed that the desulfation of HS by the Sulfs was an oriented and processive process (Seffouh et al., 2013). However, the way Sulfs recognized and bound their substrate remained unknown. In order to clarify these processes, we sought to identify protein sites involved in the HSulf-2/HS interaction, to characterize the substrate recognized by Sulfs, and to investigate the role of each domain in this binding.

2. Understanding the role of a unique PTM of HSulf-2, previously identified by our group. HSulf-2 was found to be a gagosylated protein. Interestingly, the GAG chain is anchored to the HD domain of the HSulf-2 isoform only and we hypothesized that this could affect access of the enzyme to its substrate, with consequences on its diffusion within tissues. We sought then to understand the biological function of this chain, by investigating its role *in vitro* and *in vivo* during tumor progression in a mouse model of breast carcinoma xenograft.

3. Finally, we initiated structural studies of HSulfs. Solving the structure of Sulf is a major scientific challenge. Previous crystallization assays performed on HSulf were unsuccessful. Given that the protein was pure, we speculated that this could be due to structural features of

HSulf that can hinder formation of crystals or bias its quality. In agreement with this, HSulf is a glycosylated protein and its HD domain comprises a highly disordered region of 30 amino-acids residues (based on simulation of disorder score with the iupred software). This disordered region and the glycosylations could lead to improper protein stacking within the crystal. To prevent this, we proposed here to study separately each domain of HSulf. HSulf CAT-domain (SulfΔHD) will be studied by X ray crystallography, as the CAT domain is predicted to be structured and is highly homologous to previously crystalized sulfatases. In parallel, the isolated HD domain will be studied by NMR to clarify its level of structuration and its dynamics and flexibility.

2. Dissemination

The work of this PhD project presented in this manuscript has also been the subject of 7 publications and different presentations and posters.

Publications:

- **El Masri R.**, Seffouh A., Lortat-Jacob H., Vivès RR The "in and out" of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate. Glycoconj J. 2017 Jun;34(3):285-298.
- Hijmans RS., Shrestha P., Sarpong KA., Yazdani S., **El Masri R.**, de Jong WHA., Navis G., Vivès RR., van den Born J. High sodium diet converts renal proteoglycans into pro-inflammatory mediators in rats. PLoS One. 2017 Jun 8;12(6):e0178940.
- Seffouh I., Przybylski C., Seffouh A., **El Masri R.**, Vivès RR., Gonnet F., Daniel R. Mass spectrometry analysis of the human endosulfatase Hsulf-2. Biochem Biophys Rep. 2019 Feb 7;18:100617.
- Seffouh A. *, **El Masri R. ***, Makshakova O., Gout E, Hassoun ZEO., Andrieu JP., Lortat-Jacob H., Vivès RR. Expression and purification of recombinant extracellular sulfatase HSulf-2 allows deciphering of enzyme sub-domain coordinated role for the binding and 6-O-desulfation of heparan sulfate. Cell Mol Life Sci. 2019 May;76(9):1807-1819. ***** ***Co-first authors***.
- Ferreras L., Moles A., Situmorang GR., **El Masri R.**, Wilson IL., Cooke K., Thompson E., Kusche-Gullberg M., Vivès RR., Sheerin NS., Ali S. Heparan sulfate in chronic kidney diseases: Exploring the role of 3-O-sulfation. Biochim Biophys Acta Gen Subj. 2019 May;1863(5):839-848.
- Oshima K., Yang Y., Haeger S. M., McMurtry S. A., Lane T., **El Masri R.**, Zhang F., Yue X., Vivès R.R., Linhardt R. J., Schmidt E. P. Loss of pulmonary endothelial Sulfatase-1 after experimental sepsis attenuates subsequent inflammatory responses. Am J Physiol Lung Cell Mol Physiol. 2019 Nov 1;317(5):L667-L677.
- **El Masri R**., Seffouh A., Roelants C., Gout E., Pérard J., Crétinon Y., Lortat-Jacob H., Filhol O., and Vivès RR. The sweet side of extracellular sulfatases. In preparation

Presentations & posters:

- 100th anniversary of Heparin Discovery (poster), Paris, 2016
- Journée Annuelle des Doctorants EDCSV (poster), Grenoble, 2017
- IBS scientific day (Flash presentation + poster. Best poster award), Grenoble, 2017
- 7 Lakes Proteoglycans conference (poster), Varese, 2017
- CERMAV Glycoscience Scientific day (oral presentation), Grenoble, 2017
- Second scientific days of GDR Gagosciences (poster), Grenoble, 2018
- Proteoglycans Conference GRS GRC (oral presentation + poster. Best poster award), Andover USA, 2018
- Third scientific days of GDR Gagosciences (oral presentation + poster), Lille, 2018
- Inserm Workshop on GAG biology (poster), Bordeaux, 2019

CHAPTER

III

Catalytic Mechanisms of HSulf-2

The mechanisms of HS desulfation by the Sulfs are still unclear. Recently in our lab, it has been shown that Sulfs desulfate HS by a processive and oriented process, starting from the non-reducing end of HS NS domains and moving towards the reducing end. One of the main objectives of this thesis was to study further the catalytic mechanism of the Sulfs, and notably the recognition process of HS by these enzymes. Regarding the enzyme, a wealth of evidence in the literature indicated that HD was responsible for high affinity binding to HS, and highlighted the importance of the outer region of the HD in that binding for HSulf-1 isoform (Frese et al., 2009). Regarding the substrate, it is well established that the preferred substrate of Sulfs are the trisulfated disaccharides found within the NS domains of HS chains. However, we still ignore what are the exact protein sites of Sulfs involved in HS binding, and the type or the size of HS saccharide motifs needed for productive recognition by the Sulfs.

1. Protein sites implicated in HS/HSulf-2 interaction

This work was started by Amal Seffouh, a former PhD student in the SAGAG group. To investigate the protein sites implicated in HS/HSulf-2 interaction, she first developed a protocol to produce and purify Sulfs recombinantly, and then using a cartography technique already established in the lab (Vivès et al., 2004), she identified two protein sites involved in the HSulf-2/HS interaction. My objective was then to understand the exact roles of these two sites.

Surprisingly, the two identified HS binding sites, V_{179}KEK and L_{401}KKK, are located within the CAT domain of the enzyme and not within the HD. In addition to containing the enzyme active site, these results suggested that CAT domain could therefore play a role in HSulf-2/HS interaction. To confirm this, site directed mutagenesis experiments were carried out on these sites to replace their basic residues (lysine and arginine) with alanines. Generated mutants were then expressed and their activities were analyzed. Results showed that substitution of these motifs did not affect the activity of the enzyme (which exhibited full arylsulfatase activity), neither the affinity of HS/Sulf interaction. However, they are important for the desulfation of HS. The modeling of CAT domain suggested cooperation of these two sites to elicit the enzyme endosulfatase activity (desulfation of natural substrate HS). From these data, we proposed a model for HSulf-2 mechanism: the HD domain of the enzyme recognizes and binds the substrate with high affinity, and presents it to the CAT domain. The bound substrate is then guided and finely aligned within the active site, by the VKEK and LKKK sites, to be specifically digested. This study was accompanied by complementary data about the function of CAT domain (HSulf-2 lacking the HD, HSulf2ΔHD) that I expressed as an isolated domain. This work is reported in details in the following article published in CMLS in February 2019.

Cellular and Molecular Life Sciences
https://doi.org/10.1007/s00018-019-03027-2

ORIGINAL ARTICLE

Expression and purification of recombinant extracellular sulfatase HSulf-2 allows deciphering of enzyme sub-domain coordinated role for the binding and 6-*O*-desulfation of heparan sulfate

Amal Seffouh[1] · Rana El Masri[1] · Olga Makshakova[2] · Evelyne Gout[1] · Zahra el Oula Hassoun[1] · Jean-pierre Andrieu[1] · Hugues Lortat-Jacob[1] · Romain R. Vivès[1,3]

Received: 4 December 2018 / Revised: 8 January 2019 / Accepted: 25 January 2019

Abstract

Through their ability to edit 6-*O*-sulfation pattern of Heparan sulfate (HS) polysaccharides, Sulf extracellular endosulfatases have emerged as critical regulators of many biological processes, including tumor progression. However, study of Sulfs remains extremely intricate and progress in characterizing their functional and structural features has been hampered by limited access to recombinant enzyme. In this study, we unlock this critical bottleneck, by reporting an efficient expression and purification system of recombinant HSulf-2 in mammalian HEK293 cells. This novel source of enzyme enabled us to investigate the way the enzyme domain organization dictates its functional properties. By generating mutants, we confirmed previous studies that HSulf-2 catalytic (CAT) domain was sufficient to elicit arylsulfatase activity and that its hydrophilic (HD) domain was necessary for the enzyme 6-*O*-endosulfatase activity. However, we demonstrated for the first time that high-affinity binding of HS substrates occurred through the coordinated action of both domains, and we identified and characterized 2 novel HS binding sites within the CAT domain. Altogether, our findings contribute to better understand the molecular mechanism governing HSulf-2 substrate recognition and processing. Furthermore, access to purified recombinant protein opens new perspectives for the resolution of HSulf structure and molecular features, as well as for the development of Sulf-specific inhibitors.

Keywords Glycosaminoglycan · Structure–function relationships · Extracellular matrix · Glycocalyx · Heparin

Introduction

Sulfs are sulfatases that catalyze the regioselective hydrolysis of 6-*O*-sulfate groups on heparan sulfate (HS) polysaccharides. Since their discovery in 2001 [1], accumulating evidence has highlighted Sulfs as unique members amongst sulfatases, differentiating by their extracellular localization, distinct structural organization, enzymatic activity at neutral pH, and biological role as major modulator of HS function rather than mere actor of the polysaccharide recycling metabolism. HS is a linear sulfated polysaccharide of the Glycosaminoglycan (GAG) family, which is abundantly found at the cell surface and in the extracellular matrix (ECM) of most animal tissues. It is involved in many biological processes, through its ability to bind and modulate a vast repertoire of proteins, including growth factors, cytokines and morphogens, etc. [2–4]. These large interactive properties are essentially governed by specialized saccharide regions of the polysaccharide termed S-domains.

Amal Seffouh and Rana El Masri contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00018-019-03027-2) contains supplementary material, which is available to authorized users.

✉ Romain R. Vivès
romain.vives@ibs.fr

[1] Univ. Grenoble Alpes, CNRS, CEA, IBS, 38000 Grenoble, France

[2] Kazan Institute of Biochemistry and Biophysics, FRC Kazan Scientific Center of RAS, Kazan, Russian Federation

[3] IBS, 71 Avenue des Martyrs CS 10090, 38044 Grenoble Cedex 9, France

Published online: 20 February 2019

In S-domains, the original repeated *N*-acetyl glucosamine (GlcNAc)–Glucuronic acid (GlcA) disaccharide motif is extensively modified. Glucosamines are *N*-sulfated (GlcNS), GlcA can be epimerized into iduronic acid (IdoA), and *O*-sulfation can occur at C2 of IdoA, as well as at C6 (and more rarely C3) of glucosamines. Addition and distribution of sulfate groups as well as uronic acid (UA) epimers is catalyzed by stepwise series of highly regulated enzyme reaction during HS biosynthesis (for review see [5, 6]). Sulfs contribute to further regulate the polysaccharide structure, by editing HS 6-*O*-sulfation status. These enzymes show strong specificity for highly sulfated disaccharide units, which are mostly present within the inner regions of the HS functional S-domains. Hence, although Sulf-driven desulfation is structurally subtle, it dramatically affects HS function by modulating its ability to interact with many protein ligands. Consequently, Sulfs have been associated with a number of physiopathological processes, including embryo development, tissue regeneration, cancer and neurodegenerative disease [7–9].

Sulf isoforms (Sulf-1 and Sulf-2) and orthologs share a very similar molecular organization, including 2 regions that are essential for enzyme activity: the catalytic domain (CAT) and the basic/hydrophilic domain (HD) [10–12]. Sulf CAT domain displays strong homology with other mammalian sulfatases and comprises notably strictly conserved residues involved in arylsulfatase active site, including a posttranslationally modified cysteine into an *N*-formylglycine (FGly) residue. In contrast, the HD domain is a unique feature of these enzymes and is involved in high-affinity interaction with HS substrates. This domain shows no sequence homology with any other known protein, limited secondary structure prediction and poor sequence conservation amongst Sulfs, thereby suggesting that it may confer isoform-dependent substrate-binding specificities to the enzyme. Within HS, Sulfs primarily target [UA(2S)-GlcNS(6S)] trisulfated disaccharides, although limited activity has also been reported on [UA-GlcNS(6S)] disulfated disaccharide species [11, 13, 14]. However, a wealth of evidence reported divergent activities of Sulf isoforms, notably in cancer, and studies on mouse Sulf-1 and Sulf-2 KO mice have highlighted differences in HS sulfation patterns [15–17]. Altogether, these suggest that substrate specificities of the Sulfs may not be restricted to monosaccharide or disaccharide units, but may most likely involve the recognition of much longer saccharide motifs. Recently, we have shown that human Sulfs (HSulfs) catalyzed the 6-*O*-desulfation of HS through an original, orientated and processive mechanism [13]. We speculated that this unique mechanism involved the coordinated actions of both CAT and HD domains: the HD domain would provide substrate binding and specificity, direct proper presentation to the CAT domain and drive processivity through multiple and transient interaction with HS that would allow the enzyme to "glide" along the polysaccharide chain. In agreement with this, other studies have since shown that HSulf-1 substrate recognition was complex and involved multiple interaction events with different binding sites present within the HD domain, which provided unique dynamic properties [18]. More recently, a study on HSulf-1 showed that HD/HS binding was found to exhibit atypical catch-bond type properties, with increased lifetime when subjected to external forces [19].

However, and despite growing interest, Sulfs are highly elusive enzymes, and getting structural and molecular insights into the enzymatic mechanism remains a major scientific challenge. Such studies have primarily been hampered by limited availability of recombinant enzyme. Sulfs cannot be expressed in bacteria, as the enzyme requires post-translational modifications for activity, notably *N*-glycosylations, furin cleavage maturation and formation of the catalytic FGly residue. Recovery of Sulfs from naturally expressing or transfected mammalian cells has been achieved by us and others, but only as concentrated conditioned medium preparations, as low protein yields precluded any purification attempts. In the present study, we report for the first time the preparation of purified, recombinant HSulf-2 in mammalian cells. Access to this source of enzyme enabled us to investigate further the respective roles of Sulf domains in the enzyme catalytic activities and substrate recognition process. Using heparin-bead crosslinking experiments, we identified 2 novel HS binding epitopes within the enzyme CAT domain, and we analyzed the contribution of these epitopes in the enzyme functional properties, including: (i) its arylsulfatase activity (imparted by the sulfatase-conserved active site); (ii) endosulfatase activity (Sulf-specific ability to 6-*O*-desulfate HS and heparin); (iii) and its ability to bind to heparin and HS with high affinity. Based on the data obtained, we propose here a refined model of HS 6-*O*-desulfation process by the Sulfs, which may help understanding further this complex regulatory mechanism of HS function.

Materials and methods

Unless specified otherwise, all chemicals and reagents were from Sigma.

Production of recombinant WT and mutant HSulf-2

HSulf-2 coding sequence (Genbank CR749319.1, cDNA in pcDNA3.1 plasmid was courtesy of Professor S. Rosen, University of California, USA) was amplified by PCR (for primer sequences, see Supplementary experimental 1), then inserted through EcoRV/XmaI restriction sites in a pcDNA3.1/Myc-His(-) vector modified to express proteins

of interest between TEV cleavable SNAP and 6His tags at the N- and C-terminus, respectively (gift from P. Desprès, Institut Pasteur, France). Mutants were generated from this vector by the Robiomol platform of Integrated Structural Biology Grenoble (ISBG).

WT and mutant encoding vectors were used to stably transfect FreeStyle HEK 293-F cells (medium and tissue culture reagents from Thermo fisher scientific), as previously described [13]. Protein productions were achieved by seeding cells at a 10^6/mL density in FreeStyle 293 Expression Medium, then harvesting conditioned medium (CM) 5 days later. Full-length protein purification was achieved in two steps of cation-exchange and size-exclusion chromatographies. Conditioned medium (CM) was first loaded onto a SP-Sepharose column (GE healthcare) equilibrated in 50 mM Tris, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 100 mM NaCl, pH 7.5, at a flow rate of 0.5 mL/min. After washing the column with the same buffer, proteins were eluted with a NaCl gradient (from 100 mM to 1 M). Fractions corresponding to the absorbance peak at 280 nm (0.5 mL/fraction) were collected, pooled and concentrated on a 30-kDa centrifugal unit (Centricon-30, Millipore). Concentrated samples (200 µl) were then injected onto a Superdex-200 column (GE healthcare) equilibrated with 50 mM Tris, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 300 mM NaCl, pH 7.5, at a flow rate of 0.5 mL/min. As previously, 0.5 mL fractions were collected, pooled and concentrated.

Purification of HSulf-2ΔHD (and mutants) was performed by affinity chromatography, using a nickel column (Thermo fisher scientific). CM was loaded onto a nickel resin equilibrated in 50 mM Tris buffer, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 100 mM NaCl, pH 7.5. Protein was eluted from the column with 50 mM Tris buffer, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 100 mM NaCl 350 mM imidazole, pH 7.5, then concentrated on Centricon-30, with repeated washings of the centrifugal unit with equilibration buffer to achieve complete removal of imidazole.

Purified proteins were supplemented with anti-proteases (Complete EDTA-Free, Roche), 20% glycerol, quantified and stored at − 20 °C. Detection after PAGE analysis was performed in reducing conditions, using standard protocols of Coomassie blue staining, Western-blotting with a primary goat polyclonal anti N-terminal HSulf-2 antibody (A-18, Santa-Cruz biotechnology, dil. 1/1000), or detection of the SNAP tag using SNAP fluorescent ligand SNAP-Vista Green (New England Biolabs).

Heparin-bead cross-linking experiments

Heparin-bead cross-linking analysis was performed as described previously [20]. Briefly, heparin-beads were activated with a mixture of 40 mM l-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), 10 mM *N*-hydroxysuccinimide (NHS) in 50 mM MES and 150 mM NaCl pH5.5, for 10 min at room temperature (RT). EDC/NHS in excess were inactivated by addition of ß-mercaptoethanol (15 mM final) and removed by three steps of centrifugation/washing of the beads with PBS. Heparin-beads were then incubated with 35 µg (~ 1 µM) of recombinant HSulf-2 in PBS for 2 h at RT, under gentle agitation. After quenching the reaction by addition of primary amine containing buffer (100 mM Tris, final concentration), beads were rinsed with PBS, 2 M NaCl to remove non-covalently bound material. Cross-linked HSulf-2 was denatured by heating the beads at 60 °C in PBS, 2 M Urea for 45 min and proteolyzed by incubation with thermolysine (53 mIU) at 50 °C for 16 h. Released peptides were removed by three washing steps with PBS, 2 M NaCl, 15 mM ß-mercaptoethanol, 1% Triton, while heparin-bound peptides were identified by Edman degradation automated sequencing. Cross-linked amino-acids are typically identified by the presence of a "blank" cycle during the sequencing and a drop of the recovery yields for the sequencing of downstream residues.

Aryl-sulfatase assay

Arylsulfatase activity of recombinant WT and mutant HSulf-2 was assessed using the fluorogenic pseudo-substrate 4-methyl umbelliferyl sulfate (4-MUS), as described previously {Frese, 2009 #948}. Briefly, 1–3 µg of protein was incubated with 10 mM 4-MUS in 50 mM Tris 10 mM $MgCl_2$, pH 7.5 for 1–4 h at 37 °C. Reaction was monitored by fluorescence measurement (excitation 360 nm, emission 465 nm).

Endosulfatase assay

Heparin (25 µg) was incubated with 3 µg of recombinant WT or mutant HSulf-2 in 50 µl of 50 mM Tris and 2.5 mM $MgCl_2$, pH 7.5 for 4 h. The enzyme was inactivated by heating the sample at 100 °C for 5 min, then an aliquot of the digestion products (~ 1/10) was exhaustively degraded into disaccharides by incubation with a cocktail of heparinase I, II and III (Grampian enzymes, 10 mU each) in 100 mM sodium acetate, 0.5 mM $CaCl_2$, pH 7.1 and for 48 h at 37 °C. Compositional analysis was performed by RPIP-HPLC as previously described [21]. Samples were injected onto a Phenomenex Luna 5 µm C18 reversed phase column (4.6 × 300 mm, Phenomenex) equilibrated at 0.5 mL/min in 1.2 mM tetra-*N*-butylammonium hydrogen sulfate (TBA) and 8.5% acetonitrile, then resolved using a multi-step NaCl gradient (0–30 mM in 1 min, 30–90 mM in 39 min, 90–228 mM in 2 min, 228 mM for 4 min, 228–300 mM in 2 min and 300 mM for 4 min) calibrated with HS disaccharide standards (Iduron). Post-column disaccharide derivatization was achieved by on-line addition of 2-cyanoacetamide

(0.25%) in NaOH (0.5%) at a flow rate of 0.16 mL/min, followed by fluorescence detection (excitation 346 nm, emission 410 nm).

Analysis of HSulf-2–heparin interaction by immunoassay

Binding to heparin was evaluated by an immunoassay test. Microtiter plates (black maxisorp 96wells, Nunc) were coated overnight at 4 °C with 1 mg/mL of streptavidin in 50 mM Tris–Cl and 150 mM NaCl, pH 7.5 (TBS) buffer. Plates were washed with TBS and then incubated with biotinylated heparin (1 mg/mL) prepared as described (Supplementary experimental 2), in TBS for 1 h at RT (100 µl/well). After blocking for 1 h at RT with TBS, 2% BSA, recombinant HSulf-2 variants were added to the wells at different concentrations in TBS 0.05% (w/v) tween 20 (TBS-T) and incubated at 4 °C for 2 h. Wells were washed with TBS-T and then incubated with rabbit polyclonal anti-HSulf-2 (gift from K. Uchimura, 1/1000 dilution) in TBS-T (for full-length HSulf-2), or with anti penta-HIS antibody for HSulf-2ΔHD variants (Qiagen, 1/100 dilution), at 4 °C, for 2 h or overnight, respectively. After extensive washing, fluorescent (A488) or HRP-conjugated secondary antibodies (Jackson ImmunoResearch Laboratories, 1:1000 dilution) was added for 1 h at 4 °C. Wells were washed again prior to fluorescence reading (excitation 485 nm, emission 535 nm) or treated with ECL (Thermo fisher scientific) for luminescence measurement.

Analysis of HSulf-2–heparin interaction by SPR

All experiments were performed on a BIAcore T200 (GE healthcare). Biotinylated heparin (See Supplementary experimental 2) was immobilized on a S-CM4 sensorchip, as described before [22]. Briefly, two sensorchip flow-cells were activated with a mix of 0.2 M *N*-ethyl-*N*′-(diethylaminopropyl)-carbodiimide (EDC) and 0.05 M *N*-hydroxy-succimide (NHS). Streptavidin (50 µg/mL in 10 mM acetate buffer, pH 4.5) was then injected over the activated flow cells (~ 2500 RU (Response Unit) of immobilized streptavidin). One of these flowcells served as negative control, while biotinylated heparin was injected on the other (40–50 RU of immobilized heparin). All SPR experiments were then performed, using HBS-P buffer (10 mM HEPES, 150 mM NaCl, 0.05% surfactant P20, pH 7.4) supplemented with 0.02 M EDTA), at a flow rate of 10–50 µl/min. Interaction assays involved 5-min injections of 0–40 nM HSulf-2 (WT and mutant) over the heparin and negative control surfaces, followed by a 5-min washing step with HBS-P buffer to allow dissociation of the complexes formed. At the end of each cycle, the heparin surface was regenerated by a 2.5-min injection of 2 M NaCl. Sensorgrams shown correspond to on-line subtraction of the negative control to the heparin surface signal.

Molecular modeling

For homology modeling, the sequence of HSulf-2ΔHD truncated form shown in Supplementary (Supp. Figure 2) was taken. Multiple alignment with arylsulfatase of known crystal structure (A, B and C, pdb id: 1auk, 1p49 and 1fsu, correspondingly) was performed using ClustalW algorithm to determine conservative fragments in the HSulf sequence. The HSulf-2ΔHD model was built using Modeller [23] and arylsulfatase A was chosen as a template due to its best resolution (2.1 Å). The conservative core folding was kept as in the template structure whilst the structurally variable regions including missing loops were refined. Then, the structure was repeatedly energy minimized and equilibrated in the course of short MD runs. Resulted structure was quality checked using PROCHECK online service [24].

To determine the orientation of HS towards HSulf-2ΔHD binding site, molecular docking procedure was carried out. The heparin oligosaccharide fragments GlcNS(6S)-[IdoA(2S)-GlcNS(6S)]$_n$ were used as a ligand (n amonted from 1 to 5), initial coordinates of which were taken from NMR structure of highly sulfated heparin (pdb id 1hpn). Molecular docking the HS fragments to HSulf-2ΔHD was performed using the High Ambiguity Driven biomolecular DOCKing (HADDOCK) approach [25, 26]. The algorithm allows one to perform knowledge-based docking; thus the residues determined using cross-linking mapping approach were indicated as directly involved into interaction with the ligand. The docking protocol was the following: 1000 structures were generated for initial rigid docking, then 200 of the most energetically favorable structures were subjected to semi-flexible annealing, where the parts involved in interactions were allowed to move. Afterwards, the structure of protein–ligand complexes was refined in explicit water solvent bath in the course of short molecular dynamics simulation. The most energetically favorable docking pose was taken for further analysis.

Results

Expression and purification of recombinant HSulf-2 and HSulf-2ΔHD

Purification of recombinant HSulf-2 was achieved by taking advantage of our previously reported expression system in mammalian HEK 293-F cells [13], which showed efficient protein production from high-density suspension cultures, and recovery of the secreted enzyme in low-protein, serum-free conditioned medium. For this, HSulf-2 cDNA was

inserted in a pcDNA3.1 vector modified for adding SNAP and 6His tags (at the N-ter and C-ter of HSulf-2, respectively, Fig. 1a) to facilitate protein purification and detection. We first attempted to purify the enzyme from conditioned medium by his tag affinity chromatography. However, results showed that the protein was not retained on the nickel column, even in the absence of imidazole (data not shown). We thus developed a two-step purification procedure involving cation-exchange and size-exclusion chromatographies and monitored protein elution by PAGE analysis of the collected fractions. HSulf-2 eluted from the SP-Sepharose column at ~0.4 to 0.6 M NaCL, along with other protein contaminants (Supp. Figure 1A).

Size-exclusion separation enabled recovery of pure protein (Fig. 1b, c). As reported elsewhere (Seffouh et al., submitted), only the ~95-kDa SNAP tagged N-terminal chain of the protein could be visualized by Coomassie blue staining (Fig. 1c), as confirmed by Western blotting using an anti N-terminal HSulf-2 or using a fluorescent ligand binding to the SNAP tag (Fig. 1c). However, the presence of both chains in stoichiometric abundance as well as the furin cleavage site at $R_{538}S$ was confirmed by Edman N-terminal sequencing (Supp. Figure 2). Unexpectedly, HSulf-2 eluted with a high apparent molecular weight (aMW> 1000 kDa, based on elution time), but protein aggregation and/or oligomerization was ruled out by quality control analysis of the preparation using negative staining electron microscopy (data not shown). From our data, we estimated net production yields of recombinant HSulf-2 at ~2 to 3 mg/L of culture medium and purity at 90–95%. HSulf-2 signal was also occasionally found in a second, minor, late eluting peak ($_{deg}$HSulf-2, Fig. 1b), which corresponded to degraded forms of the enzyme. As the 95-kDa band could still be visualized on PAGE analysis of the corresponding fractions, this could suggest higher sensitivity of HSulf-2 HD domain to proteolytic degradation. However, work at 4 °C and addition of antiproteases throughout the purification procedure considerably reduced the size of this second peak.

Expression of HSulf-2ΔHD in HEK293F cells was achieved similarly, but surprisingly, the protein bound

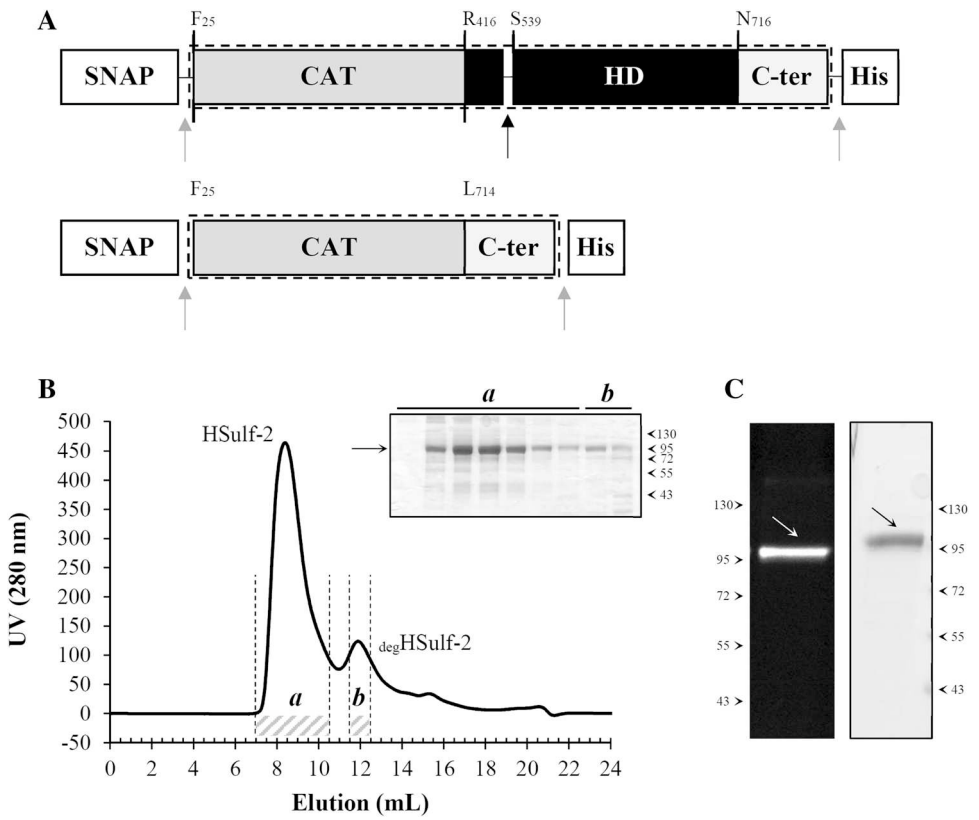


Fig. 1 Expression and purification of recombinant HSulf-2. **a** Representation of HSulf-2 and HSulf-2ΔHD constructs. Boxes represent HSulf-2 C-ter, CAT and HD domains (from light grey to black, respectively) and added SNAP and His tags (white boxes). Arrows indicate furin (black) and added-TEV (grey) cleavage sites. Amino acids (numbered according to HSulf-2 full amino acid sequence) delineating HSulf-2 domains are shown. **b** Size-exclusion separation profile of HSulf-2. Fractions corresponding to HSulf-2 (***a***) and $_{deg}$H-Sulf-2 (***b***) are indicated by grey/white dashed area and PAGE stained by Coomassie blue of these fractions are shown in the inset. The arrows indicate the band corresponding to HSulf-2 N-terminal chain. **c** Gel electrophoresis of HSulf-2 revealed by fluorescent SNAP substrate (left) and Coomassie blue (right). Arrows indicate the band corresponding to the 95 kDa SNAP-tagged N-terminal chain

efficiently to the nickel column (Supp. Figure 1B and 1C). PAGE analysis showed a single ~ 110-kDa band corresponding to the SNAP-tagged HSulf-2ΔHD (note that HSulf-2ΔHD is composed of a single polypeptide chain, as the furin cleavage site present in the HD domain is absent). An estimated ~ 90% purity was achieved using a single step His-tag affinity chromatography and production yield was significantly higher that for the Full-length enzyme (~ 8 to 10 mg/L).

Arylsulfatase, heparin binding and 6-*O*-endosulfatase activities of HSulf-2 and HSulf-2ΔHD

To validate our expression system, we next assessed the functional properties of purified recombinant HSulf-2, using three distinct biochemical assays. We first analyzed the arylsulfatase activity of recombinant HSulf-2 using the 4MUS assay, which measures the ability of arylsulfatases to convert the non-fluorescent 4-MUS pseudo-substrate into the fluorescent 4-MU product [10]. Results confirmed the ability of both enzymes to process the 4MUS, with comparable levels of fluorescence monitored after 1 h of incubation (Fig. 2a), thereby indicating that deletion of the HD domain did not compromise the integrity of the enzyme active site. Noteworthy, although both HSulf-2 and HSulf-2ΔHD catalyzed 4-MUS desulfation at very similar initial velocities, processing rate was sustained for up to 4 h for HSulf-2, but decreased after 2 h for HSulf-2ΔHD (Supp. Figure 3A).

HS binding properties of HSulf-2 and HSulf-2ΔHD were next investigated by SPR, as previously described [22]. In this assay, reducing-end biotinylated heparin (see supplementary experimental 2) is immobilized on streptavidin-coated sensorchips, a design that mimics to a certain extent the display of proteoglycan-bound HS chains at the cell surface. HSulf-2 was injected onto these functionalized surfaces, in an EDTA-supplemented buffer to prevent enzymatic degradation of the immobilized GAGs, and SPR sensorgrams of the interaction were recorded in real time. Results obtained for full length showed a very productive binding to heparin, with a slow dissociation phase suggesting formation of stable enzyme/substrate complexes (Fig. 3a). Unexpectedly, HS binding properties of HSulf-2ΔHD were not totally abolished, as lower but significant interaction to the heparin surface could still be monitored by SPR (Fig. 3a). To determine K_D and kinetic parameters for the interaction, we injected a series of HSulf-2 concentrations over the heparin surface (Fig. 3b). However, examination of the binding curves clearly indicated a complex mode of interaction and the data could not be fitted to a 1:1 Langmuir binding model. We thus analyzed the interaction using ELISA assays. Results yielded typical binding curves (Figs. 3c, d) and K_Ds were determined by Scatchard analysis. Affinities of 4.2 ± 1.2 nM and 20.4 ± 2.1 nM were calculated for full-length HSulf-2 and HSulf-2ΔHD, respectively. These results confirm the major role of the HD domain in the binding to HS, but indicate that the CAT domain on its own can bind, although with significantly lower efficiency, to HS and heparin.

We finally assessed HS 6-*O*-endosulfatase properties of the enzymes. For this, we analyzed heparin disaccharide composition, following treatment by either HSulf-2 or HSulf-2ΔHD. Data (Fig. 2b) showed that digestion with the full-length enzyme dramatically reduced heparin [ΔHexA(2S)-GlcNS(6S)] trisulfated disaccharide content (− 19% compared to untreated heparin) and concomitantly increased the level of the resulting [ΔHexA(2S)-GlcNS] disulfated disaccharide (+ 14% compared to untreated heparin). Additionally, a small but significant decrease in [ΔHexA-GlcNS(6S)] disaccharide could be observed in HSulf-2 treated heparin (− 1.4%). In contrast, HSulf-2ΔHD showed impaired 6-*O*-endosulfatase activity, with no significant changes in [ΔHexA(2S)-GlcNS(6S)] and [ΔHexA(2S)-GlcNS] composition when compared to untreated heparin.

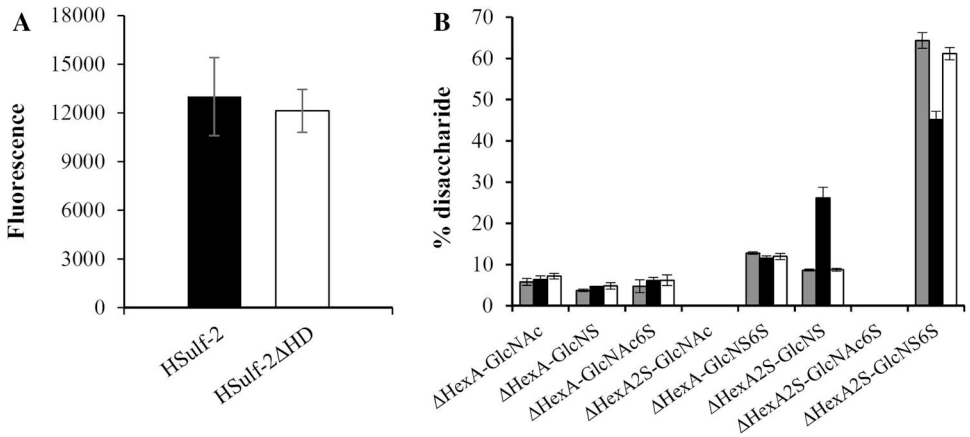


Fig. 2 Comparison of HSulf-2 and HSulf-2ΔHD arylsulfatase and endosulfatase activities. **a** Processing of 4-MUS after a 1 h digestion with HSulf-2 (black) and HSulf-2ΔHD (white). **b** Disaccharide analysis of heparin, without (grey) or after a 4 h digestion with HSulf-2 (black) or HSulf-2ΔHD (white). Error bars represent SEM of triplicate analysis

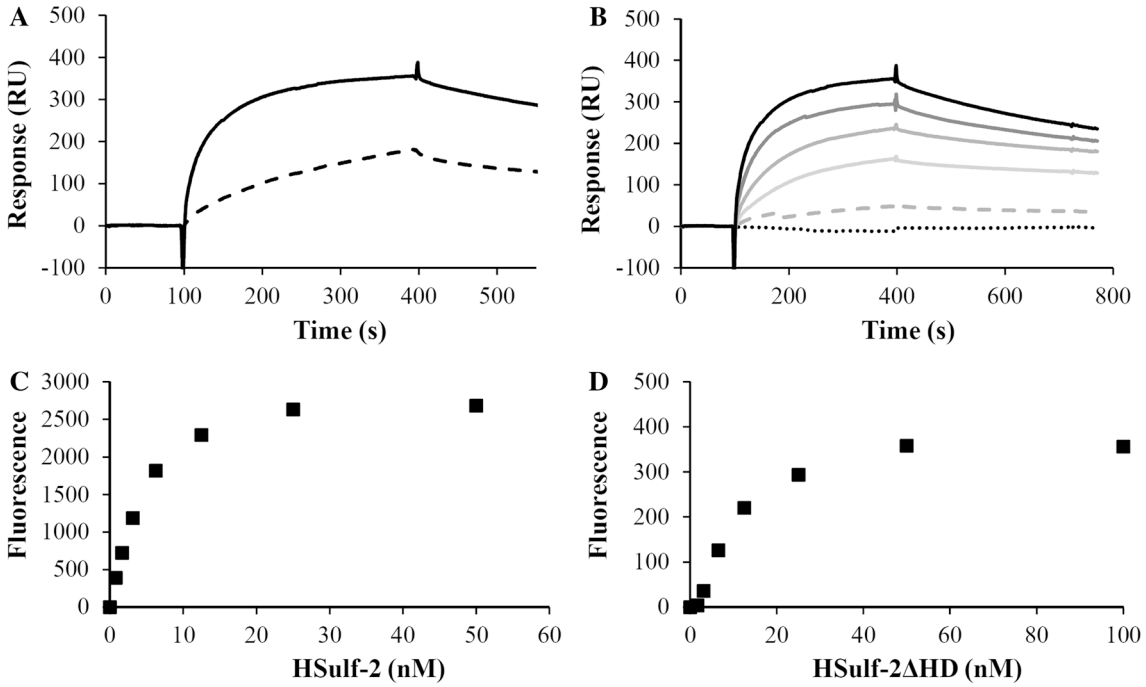


Fig. 3 Interaction of HSulf-2 and HSulf-2ΔHD with heparin. **a** SPR analysis of the binding of HSulf-2 (plain) and HSulf-2ΔHD (dashed) on a heparin-functionalized surface. **b** Injection of buffer (dashed, black line) or increasing (from dashed pale, to plain dark) 1.25–40 nM concentrations of HSulf-2 on the heparin surface. **c** Immunoassay of HSulf-2 interaction with heparin. **d** Immunoassay of HSulf-2ΔHD interaction with heparin; curves shown are representative of at least three independent experiments

Altogether, these results corroborate previously reported HSulf-2 activities, thereby confirming the functional integrity of our purified, recombinant enzyme.

Mapping of HSulf-2 HS binding sites

The development of an efficient recombinant HSulf-2 production system is an important step forward that opens wide perspectives for studying the structural and functional features of this enzyme. Here, we chose to focus on the characterization of the enzyme/substrate recognition process. We thus used an in-house developed technique to cartography heparin binding sites within proteins [20], with the aim to identify HSulf-2 amino acids involved in HS binding. This method is based on the formation of covalent complexes between the protein of interest and heparin functionalized beads, the proteolytic digestion of these complexes and the identification of the peptides remaining trapped on the polysaccharide (i.e. participating to the binding) by performing N-terminal sequencing analysis directly on the beads.

Analysis of HSulf-2 using this approach yielded data with unusually high background (i.e. amino acids that could not be confidently attributed to any sequence within HSulf-2) and a rapid drop of recovery yields at each sequencing cycle. Peptides located within the HD domain could be detected, but these results were erratic and poorly reliable. In contrast, the approach consistently highlighted two short peptides within the enzyme CAT domain: V_{179}KEK and L_{401}KKK (Supp. Figure 2).

Generation and functional characterization of HSulf-2 mutants

To investigate the functional relevance of the two putative HS binding sites identified using the cross-linking approach, we generated enzyme mutants, in which lysine residues of the VKEK and LKKK motifs were replaced by alanines. Alanine substitutions were introduced by site-directed mutagenesis to generate plasmids encoding two single motif mutants HSulf-2/VAEA, HSulf-2/LAAA and the double-mutant HSulf-2/VAEA/LAAA. As for HSulf-2 WT, vectors were used to stably transfect HEK293 cells. Enzymes were then recovered from the conditioned medium and purified as described above, with no changes in the expression yields and purification procedure.

We first sought to determine whether these mutations affected the enzyme arylsulfatase activity, using the 4MUS assay. Results (Fig. 4a and Supp. Figure 3B) showed equivalent fluorescence signal levels for WT and all mutant enzymes, indicating that neither VKEK nor LKKK epitopes are involved in HSulf arylsulfatase activity. We next compared the ability of WT and mutant HSulf-2 to bind to

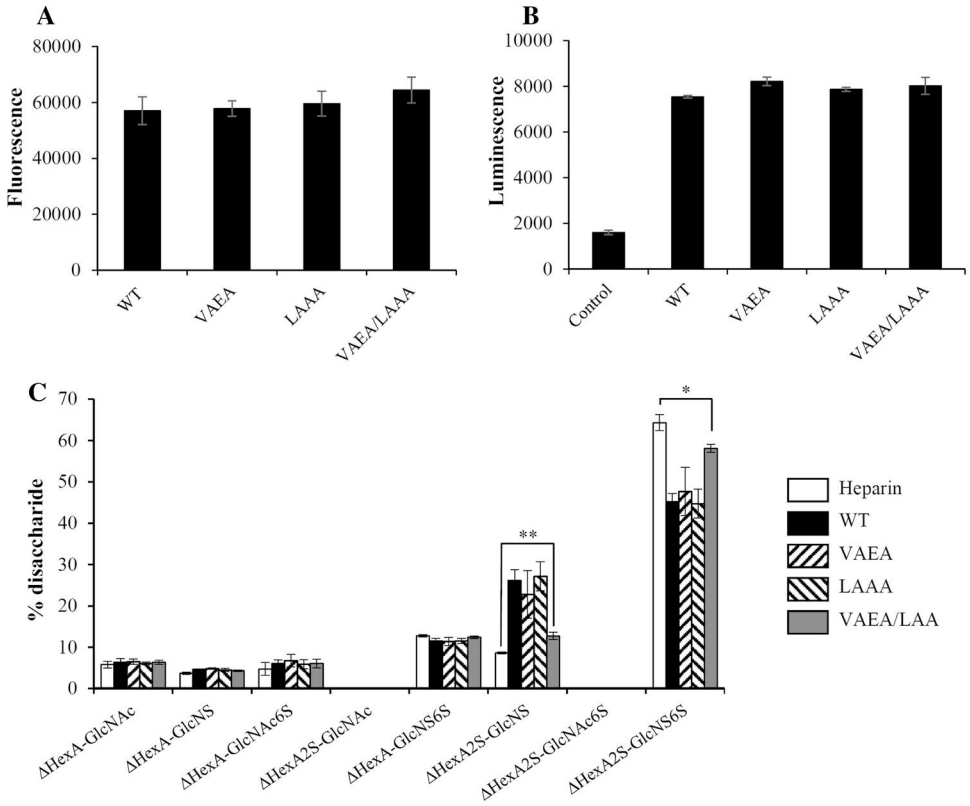


Fig. 4 Comparison of HSulf-2 and HSulf-2 mutants' biological activities. Analysis of the 4-MUS (**a**), heparin binding (**b**) and 6-*O*-endosulfatase (**c**) activities of WT HSulf-2 and mutants. Error bars represent SEM of triplicate analysis

heparin in our immunoassay. Results showed that all forms tested effectively bound to the polysaccharide (Fig. 4b). A K_D of 4.7 ± 3.6 nM was determined for the interaction of HSulf-2/VAEA/LAAA to heparin (Supp. Figure 4), an affinity very similar to that exhibited by HSulf-2 WT (Fig. 3c). Noteworthy, introduction of the VAEA/LAAA mutations in the HSulf-2ΔHD truncated form significantly reduced, but did not completely abolish binding (Supp. Figure 4). Altogether, these data suggest that HSulf-2 VKEK and LKKK epitopes are primarily responsible for the low-affinity HS binding property of the enzyme CAT domain (although other residues within this domain may also be involved), but contribution of these motifs to the interaction of the full-length enzyme with the polysaccharide is negligible compared to that of the HD domain.

HS 6-*O*-endosulfatase activity was finally assessed, as described above (Fig. 4c). When compared to HSulf-2 WT, treatment with HSulf-2/VAEA and HSulf-2/LAAA yielded very similar composition profile. In contrast, HSulf-2/VAEA/LAAA double mutant led to drastically reduced changes in heparin disaccharide composition, with attenuated but significant 6.2% reduction and 4.1% increase in [ΔHexA(2S)-GlcNS(6S)] and [ΔHexA(2S)-GlcNS] disaccharide content, respectively. Importantly, these results indicate that combined mutations of VKEK and LKKK epitopes significantly reduce, but do not abrogate the enzyme 6-*O*-endosulfatase activity. In agreement with this, we found that increased 6-*O*-desulfation of heparin by HSulf-2/VAEA/LAAA could be achieved using more extensive digestion conditions (data not shown).

Conservation of the CAT domain HS-binding sites within sulfatases

In light of these observations, we then studied the conservation of these newly identified functional epitopes amongst sulfatases (Fig. 5). Despite significant homology of Sulf CAT domain with other sulfatases, VKEK and LKKK like sequences were only found in Sulf isoforms and orthologs, thus supporting further a contribution to the specific HS 6-*O*-endosulfatase activity of these enzymes. Noteworthy, HSulf-2 K_{402} of the L_{401}KKK motif was only conserved in Sulf-2 orthologs and replaced by an asparagine in Sulf-1. This could either suggest that this residue could be involved in Sulf-2 isoform-specific processing of HS substrate, or may not play any relevant role. Similarly, residue at position 179 of the V_{179}KEK motif was consistently a non-polar amino acid in mammalian Sulfs (V for Sulf-2 or I for Sulf-1), but was substituted by an asparagine in Quail Sulf-1 (QSulf-1). Finally, it is worth noting that Glucosamine 6-sulfatase exhibits three basic residues aligning with HSulf-2 VKEK motif, which may suggest that this epitope could partly participate to the specific recognition of glucosamine 6-sulfate residues.

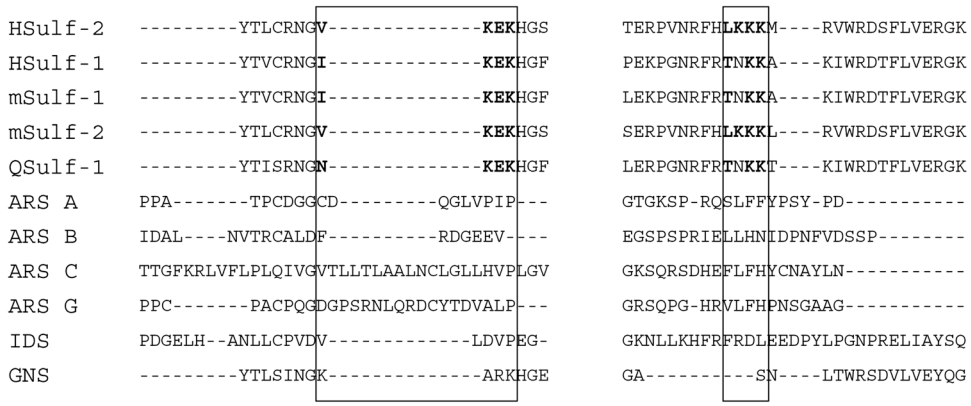


Fig. 5 Conservation of CAT HS binding epitopes within sulfatases. Clustal omega sequence alignment of HSulf-2 and HSulf-1; mouse Sulf-1 and -2 (mSulf-1 and mSulf-2, respectively); quail Sulf-1 (QSulf-1); human arylsulfatase (ARS) A, B, C and G; human iduronate 2-sulfatase (IDS) and human Glucosamine 6-sulfatase (GNS). Sequences aligned with HSulf-2 VKEK and LKKK epitopes are framed and conserved residues within these motifs are in bold

Molecular modelling of heparin–HSulf-2 Cat-domain interaction

To provide a structural basis for the underlying mechanism, we used HSulf-2 CAT domain homology with sulfatase of known structure to model the interaction of the VKEK and LKKK epitopes with a heparin oligosaccharide. The spatial structure of HSulf-2ΔHD was built on the base of three human arylsulfatases alignment. The sequences of three arylsulfatases share 29–36% of identity and demonstrate common folding. CAT-domain of Sulf2 reveals 26–28% of identity toward these arylsulfatases. Resulted model structure was quality controlled with following output (corresponding values for the template structure are given in parenthesis): G-factors overall −0.26 (−0.14), covalent −0.41 (−0.22), dihedrals −0.18 (−0.15), torsions in favorable regions were 82.6% (86.8), in allowed 14% (11.9), in generously allowed 2.6% (1.1) and in disallowed regions 0.9% (0.3). The model indicated the alignment of L_{401}KKK, FGly and V_{179}KEK within a groove of the protein surface, forming a binding site (Fig. 6a). The distance between the fragments L_{401}KKK and V_{179}KEK, calculated for Cα atoms belonging to K_{404} and K_{180}, amounts to 32 Å and could thus accommodate an octasaccharide (Fig. 6b). Given the 6-sulfate of the GlcNS(6S) in +1 position facing the catalytic FGly, the GlcNS(6S) in −4 position is in front of K_{404}, thereby establishing an electrostatic contact between the 6-sulfate and the lysine side chain amino group. Furthermore, the IdoA(2S) in +3 position occurs next to K_{180} forming a salt bridge between the 2-sulfate group and the amino group of lysine. A longer oligosaccharide, GlcNS(6S)-[IdoA(2S)-GlcNS(6S)]$_5$, is required to grasp both lysines in V_{179}KEK fragment of the binding site. Herein, an additional salt bridge is formed between the 2-sulfate of the GlcNS(6S) in +6 position and the amino group of the K_{182}.

Discussion

Whereas all other sulfatases take part in the cellular catabolism of sulfated compounds, enzymes of the Sulf sub-family have been associated with major regulatory processes. This

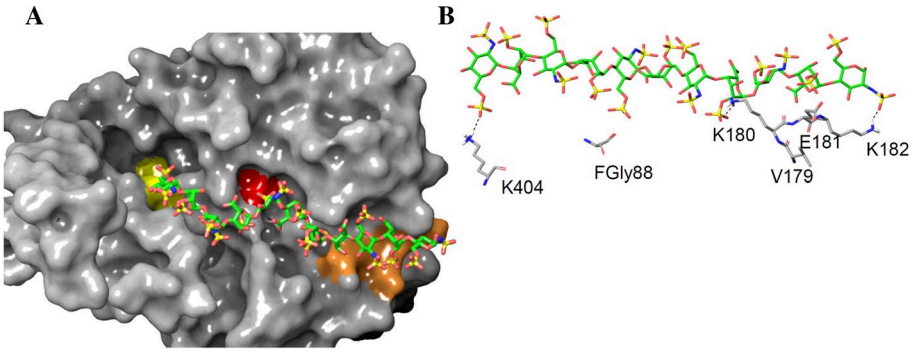


Fig. 6 Modeling of the HSulf-2ΔHD in complex with a heparin oligosaccharide. The surface representation of HSulf-2ΔHD model with the ligand GlcNS(6S)-[IdoA(2S)-GlcNS(6S)]$_5$ bound. Catalytic cysteine modified FGly (C88) is highlighted in red and the residues of V_{179}KEK and L_{401}KKK sequences identified using the cross-linking mapping approach are shown in orange and yellow, respectively (**a**). View of protein–HS ligand complex showing the minimal length unit required for efficient binding of two epitopes. Color coding is the following: C atoms—grey for the protein and green for the ligand; N atoms—blue, O atoms—red, S atoms—yellow, H atoms (given only for lysine amino groups)—white (**b**)

functional specificity is both due to the complex nature of their HS polysaccharide substrate, which binds and modulates a multitude of protein ligands, and to their endosulfatase nature (other sulfatases are exoenzymes) that confers them ability to target and edit 6-*O*-sulfation pattern of HS functional S-domains. In addition, we have recently shown that HSulfs catalyzed HS 6-*O*-desulfation, through an original, processive and orientated mechanism [13]. All these unique features clearly suggest a highly complex mode of enzyme–substrate recognition, which largely remains to be clarified. However, mechanistic and structural studies of Sulfs remain scarce and have been limited to the difficulty to achieve preparation of pure recombinant and functional enzymes. We report here for the first time the preparation of purified, recombinant HSulf-2, which was achieved by expressing the enzyme in mammalian HEK293F cells adapted to culture in suspension. Such system facilitated purification of HSulf-2 from a conditioned medium depleted in protein contaminants (HEK293F cells are grown in serum-free medium), and guaranteed the integrity of Hsulf post-translational modifications, as demonstrated by mass spectrometry proteomic analysis (Seffouh et al., submitted). Surprisingly, attempts to purify the protein through its 6xHis C-terminal tag were unsuccessful, suggesting that the C-Terminal end of the protein may not be available within the folded protein. Nevertheless, efficient purification could be achieved using standard ion-exchange and size exclusion chromatography procedures. Noteworthy, yields for the preparation of full-length purified (90–95% purity) HSulf-2 reached ~ 2 to 3 mg/L of culture medium, a figure that now opens new perspectives of studying HSulf-2 through material demanding biophysical and structural analysis techniques. Intriguingly, HSulf-2 full-length protein eluted with a high aMW, while the HSulf-2ΔHD mutant (i.e. HSulf-2 lacking its HD domain) showed size-exclusion elution time in accordance with predicted aMW (data not shown). In addition, the C-terminal chain of the enzyme including most of the HD domain could not be visualized by Coomassie blue-stained PAGE (Fig. 1c). These observations along with our proteomic analysis data (Seffouh et al., submitted) clearly suggest unusual structural properties within the HD domain. We speculate that this domain could feature an unusually extended conformation and/or multiple *N*-glycosylation/post-translational modifications, which would significantly increase the hydrodynamic radius of the whole protein. In addition, the detection of HSulf-2 degradation products, most likely within the HD domain, suggests propensity to proteolysis (contrary to the CAT domain), in line with secondary structure predictions based on amino-acid sequence, which indicated the presence of unstructured regions within the HD domain. Further investigations will be needed to clarify HD unexpected structural and functional behavior.

Consistently with previous studies performed on concentrated CM [10, 11, 13], the purified recombinant HSulf-2 showed all expected biological activities. We first demonstrated that the enzyme retained the ability to process the fluorogenic 4MUS pseudo-substrate, which is the hallmark of arylsulfatase activity (Fig. 2a). We also confirmed that HSulf-2 productively bound to heparin (Fig. 3). SPR analysis of HSulf-2 interaction with heparin (Fig. 3a, b) showed formation of very stable enzyme/polysaccharide complexes (limited loss of signal during the dissociation phase). Noteworthy, data could not be fitted to a Langmuir binding model. This evidenced a complex mode of interaction, in agreement with previous data suggesting the contribution of multiple and dynamic binding events within Sulf HD/heparin interaction [19]. As dynamic binding measurements failed to provide reliable kinetic parameters, binding K_D was determined by ELISA, which confirmed the expected high affinity (~ 4 nM) of HSulf-2 for heparin (Fig. 3c). Finally, we demonstrated the ability of recombinant, purified HSulf-2 to 6-*O*-desulfate heparin (Fig. 2b). In agreement with published data obtained with concentrated CM, the purified full-length enzyme essentially targeted the heparin [ΔHexA(2S)-GlcNS(6S)] trisulfated disaccharide, although residual activity could also be observed on [ΔHexA-GlcNS(6S)] disaccharides. Noteworthy, removal of tags upon TEV treatment had no effect on HSulf-2 activity (data not shown).

In contrast with full-length HSulf-2, the HSulf-2ΔHD form was efficiently retained by nickel column chromatography (Supp. Figure 1B), thus suggesting a contribution of the HD domain in burying the C-terminal domain within the native structure. Yields obtained were also significantly higher. We assume that this could be due to a reduced retention of the enzyme by cell-surface HS, in agreement with studies highlighting the HD domain as the major cell-surface binding component [11]. HSulf-2ΔHD exhibited similar 4MUS activity (Fig. 2a), but with a shorter initial velocity period (Supp. Figure 3A). This discrepancy should be investigated further and may indicate a role of the HD domain in the stability of the enzyme. In contrast, HSulf-2ΔHD exhibited strongly reduced heparin binding (Fig. 3) and 6-*O*-endosulfatase activity (Fig. 2b). This is consistent with the critical role played by the HD domain in the interaction with HS substrate. However, it is worth noting that residual binding and activity could still be noted, thus suggesting that the HD domain significantly contributes but may not be strictly required for the enzyme 6-*O*-endosulfatase activity and that additional HS binding sites may be found within the enzyme CAT domain.

Altogether, these data validated the functional integrity of recombinant purified HSulf-2 and ruled out the possibility of additional partners possibly present in concentrated medium preparations, which could play a role in the binding and/or processing of HS. We also further confirmed

previous indications that HSulf-2 CAT and C-ter domains (i.e. HSulf-2ΔHD) comprised all necessary features required for a functional arylsulfatase active site [12], and that Sulf HD domain plays a major role in high-affinity binding to HS heparin [11, 12, 27].

We next took advantage of this newly available source of purified HSulf-2 to clarify the molecular determinants involved in the enzyme–substrate recognition process. Because of the abundance of basic residues within Sulf amino acid sequence (127 K and R residues for HSulf-2), identification of HS binding domains could not readily be investigated through conventional alanine scanning site directed mutagenesis approaches. We thus addressed this issue, using a previously developed cross-linking technique to map HS binding sites [20]. Briefly, this technique is based on the generation of protein/heparin bead covalent complexes, the proteolytic digestion of these complexes and the identification of peptides remaining bound to the heparin (i.e. comprising the heparin binding site) by N-terminal sequencing. A first examination of HD highly basic 299 amino acid sequence revealed the presence of 2 putative Cardin-Weintraub HS binding motifs [28]: R_{518}RKKLFKKKYK and K_{703}RKKKLRKLLKR. However and although this technique has been successfully applied to the study of many different heparin binding proteins over the years, analysis of HSulf-2 yielded unexpected results. The first cycles of sequencing showed high background level and recovery yields dropped very rapidly. Consequently, no HS binding sequence could be clearly identified within the HD domain. A likely explanation of this is the presence of too many different peptides cross-linked to the heparin to allow unambiguous amino acid attribution. We thus speculate that heparin covalently bound to many amongst the HD 41 NHS-activable lysine residues and that this interaction may not simply involve the two putative Cardin–Weintraub HS binding motifs. This again is in agreement with studies, showing that interaction of HS with HSulf HD domain was highly dynamic and involved multiple binding sites [19, 29]. Furthermore, generation of very short peptides also supports further a high susceptibility of this domain to enzyme proteolysis and the presence of large poorly structured regions within the HD domain. Nevertheless, our cross-linking strategy enabled identification of 2 epitopes present within the enzyme CAT domain: V_{179}KEK and L_{401}KKK. We thus generated mutants lacking each (or both) of these domains and analyzed their enzymatic activity. Data obtained showed no difference between HSulf-2 wild-type and mutant forms in the 4MUS assay (Fig. 4a and Supp. Figure 3B), suggesting that VKEK and LKKK epitopes were not part of the enzyme active site per se. Binding to heparin was not compromised either, which was expected as the HD domain is the major contributor to this interaction (Fig. 4b). Finally, we found very similar 6-*O*-endosulfatase activities for wild-type and single mutants, whereas HSulf-2 VAEA/LAAA double mutant exhibited dramatically reduced ability to 6-*O*-desulfate heparin (Fig. 4c). These results thus suggest that VKEK/LKKK epitopes are not required for the endosulfatase activity, but cooperatively contribute to its efficiency. These experimental data were reinforced by analysis of sulfatase sequences, which showed that these motifs were only conserved in Sulf orthologs (Fig. 5).

We next modelled HSulf-2ΔHD in complex with a GlcNS(6S)-[IdoA(2S)-GlcNS(6S)]$_5$ heparin oligosaccharide. The validity of such model is strongly supported by the sequence homology between HSulf-2 CAT domain and structure-solved arylsulfatases (A, B and C), which all share a conserved protein core folding. In absence of structural data, it should nevertheless be pointed out that the presence of the HD domain could impact the CAT domain structure in the context of the whole protein. However, we can speculate that a partially unstructured HD would have minor effects over the structure of the well-folded CAT domain. Our model showed alignment of the VKEK and LKKK epitopes with active site FGly residue, and that an octasaccharide could span over all three motifs. As model calculations were made on the basis of rigid proteins, we anticipate that potential local flexibility could slightly affect the distance between the VKEK and LKKK motifs. Precise determination of the oligosaccharide size will thus require further experimental evidence. However, this suggests that VKEK and LKKK binding sites participate to the endosulfatase activity by enabling recognition and binding of larger HS motifs than the [IdoA(2S)-GlcNS(6S)] disaccharide substrate (Fig. 6). Noteworthy, non-conserved K_{402} of the L_{401}KKK motif did not establish contact with the oligosaccharide, which supports the hypothesis that this residue is not involved in the binding.

On the basis of these data, we refined our previously published model [13] to explicit HSulf desulfation mechanism (Fig. 7). In this model, HSulf primary high-affinity interaction occurs through the enzyme HD domain. However, VKEK and LKKK epitopes contribute further to this interaction by guiding and properly aligning the polysaccharide towards the enzyme active site (Fig. 7a). In the absence of these motifs, productive interaction still occurs, but desulfation is inefficient due to improper presentation of the polysaccharide to the active site (Fig. 7b). Processing of small pseudo-substrates such as 4-MUS are unaffected by the removal of VKEK/LKKK motifs, as these access directly to the active site. In absence of the HD, affinity of the enzyme for HS is significantly reduced and impairs subsequent 6-*O*-desulfation (Fig. 7c).

In conclusion, we describe here a novel, robust and efficient system to produce functional, purified recombinant Sulfs. This new source of enzyme enabled us to investigate and provide further insights into the structural basis

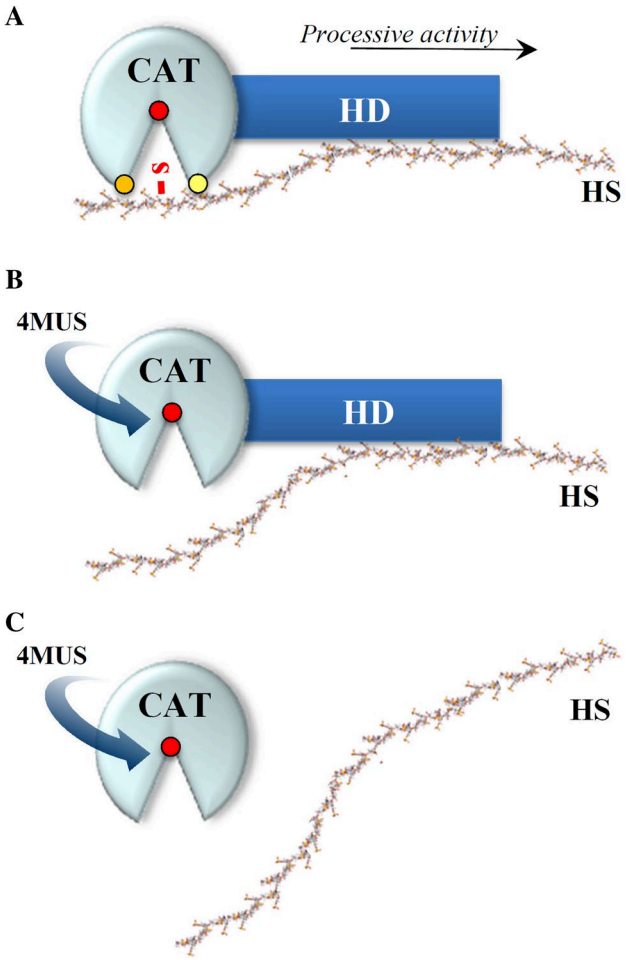


Fig. 7 Coordinated activity of CAT and HD domains for the binding and 6-*O*-desulfation of HS. For WT HSulf-2 (**a**), high-affinity binding to HS is mediated by the HD domain, while CAT domain VKEK (orange dot)/LKKK (yellow dot) HS binding sites contribute to efficient 6-*O*-desulfation by enabling presentation of the polysaccharide chain to the active site (FGly residue shown as a red dot). For HSulf-2/VAEA/LAAA mutant (**b**), high-affinity binding to HS still takes place, but the polysaccharide chain is not properly guided towards the active site, leading to impaired 6-*O*-endosulfatase activity. Processing of small 4-MUS pseudo-substrate remains unaffected. Further removal of the HD domain (**c**) leads to a HSulf-2ΔHD/VAEA/LAAA mutant unable to bind or process HS substrates

of HSulf-2 enzyme/substrate recognition. Based on our data, we propose here a refined model of the underlying mechanism that highlights the coordinated functions of Sulf CAT and HD domains, which constitutes a significant step forward towards the understanding of the highly complex and elusive enzymes. Finally, we foresee that access to purified recombinant Sulf with expression yield in the mg/L range should pave the way to new developments and major progress in the characterization of Sulf structure and molecular features, as well as for the screening and design of Sulf-specific inhibitors for potential application in cancer therapy.

Acknowledgements The authors would like to thank Elisa Tournebize for technical assistance, Marjolaine Noirclerc-Savoye for her precious advice on molecular biology, Philippe Desprès for providing the SNAP-containing shuttle vector and Kenji Uchimura for the Anti-HSulf-2 antibody. This work used the SPR, Robiomol and amino-acid sequencing platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). This work was also supported by the CNRS and the GDR GAG (GDR 3739), the "Investissements d'avenir" program Glyco@Alps (ANR-15-IDEX-02), and by grants from the Agence Nationale de la Recherche (ANR-12-BSV8-0023 and ANR-17-CE11-0040) and Université Grenoble-Alpes (UGA AGIR program).

References

1. Dhoot GK, Gustafsson MK, Ai X, Sun W, Standiford DM, Emerson CP Jr (2001) Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. Science 293:1663–1666
2. Sarrazin S, Lamanna WC, Esko JD (2011) Heparan sulfate proteoglycans. Cold Spring Harb Perspect Biol. https://doi.org/10.1101/cshperspect.a004952
3. Kjellén L, Lindahl U (2018) Specificity of glycosaminoglycan-protein interactions. Curr Opin Struct Biol 50:101–108. https://doi.org/10.1016/j.sbi.2017.12.011
4. Monneau Y, Arenzana-Seisdedos F, Lortat-Jacob H (2016) The sweet spot: how GAGs help chemokines guide migrating cells. J Leukoc Biol 99:935–953. https://doi.org/10.1189/jlb.3MR0915-440R
5. Li J-P, Kusche-Gullberg M (2016) Heparan sulfate: biosynthesis, structure, and function. Int Rev Cell Mol Biol 325:215–273. https://doi.org/10.1016/bs.ircmb.2016.02.009
6. Kreuger J, Kjellen L (2012) Heparan sulfate biosynthesis: regulation and variability. J Histochem Cytochem 60:898–907. https://doi.org/10.1369/0022155412464972
7. Vives RR, Seffouh A, Lortat-Jacob H (2014) Post-synthetic regulation of HS structure: the Yin and Yang of the sulfs in cancer. Front Oncol 3:331. https://doi.org/10.3389/fonc.2013.00331
8. Rosen SD, Lemjabbar-Alaoui H (2010) Sulf-2: an extracellular modulator of cell signaling and a cancer target candidate. Expert Opin Ther Targets 14:935–949. https://doi.org/10.1517/14728222.2010.504718
9. Nishitsuji K (2018) Heparan sulfate S-domains and extracellular sulfatases (Sulfs): their possible roles in protein aggregation diseases. Glycoconj J. https://doi.org/10.1007/s10719-018-9833-8
10. 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
11. Frese MA, Milz F, Dick M, Lamanna WC, Dierks T (2009) Characterization of the human sulfatase Sulf1 and its high affinity heparin/heparan sulfate interaction domain. J Biol Chem 284:28033–28044
12. Tang R, Rosen SD (2009) Functional consequences of the subdomain organization of the sulfs. J Biol Chem 284:21505–21514
13. Seffouh A, Milz F, Przybylski C, Laguri C, Oosterhof A, Bourcier S, Sadir R, Dutkowski E, Daniel R, van Kuppevelt TH, Dierks T, Lortat-Jacob H, Vives RR (2013) HSulf sulfatases catalyze processive and oriented 6-*O*-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity. Faseb J 27:2431–2439. https://doi.org/10.1096/fj.12-226373
14. Pempe EH, Burch TC, Law CJ, Liu J (2012) Substrate specificity of 6-*O*-endosulfatase (Sulf-2) and its implications in synthesizing

anticoagulant heparan sulfate. Glycobiology 22:1353–1362. https://doi.org/10.1093/glycob/cws092
15. Nagamine S, Tamba M, Ishimine H, Araki K, Shiomi K, Okada T, Ohto T, Kunita S, Takahashi S, Wismans RG, van Kuppevelt TH, Masu M, Keino-Masu K (2012) Organ-specific sulfation patterns of heparan sulfate generated by extracellular sulfatases Sulf1 and Sulf2 in mice. J Biol Chem 287:9579–9590. https://doi.org/10.1074/jbc.M111.290262
16. Lamanna WC, Baldwin RJ, Padva M, Kalus I, Ten Dam G, van Kuppevelt TH, Gallagher JT, von Figura K, Dierks T, Merry CL (2006) Heparan sulfate 6-*O*-endosulfatases: discrete in vivo activities and functional co-operativity. Biochem J. 400:63–73
17. Lamanna WC, Frese MA, Balleininger M, Dierks T (2008) Sulf loss influences *N*-, 2-*O*-, and 6-*O*-sulfation of multiple heparan sulfate proteoglycans and modulates fibroblast growth factor signaling. J Biol Chem 283:27724–27735
18. Milz F, Harder A, Neuhaus P, Breitkreuz-Korff O, Walhorn V, Lubke T, Anselmetti D, Dierks T (2013) Cooperation of binding sites at the hydrophilic domain of cell-surface sulfatase Sulf1 allows for dynamic interaction of the enzyme with its substrate heparan sulfate. Biochim Biophys Acta. https://doi.org/10.1016/j.bbagen.2013.07.014
19. Harder A, Möller A-K, Milz F, Neuhaus P, Walhorn V, Dierks T, Anselmetti D (2015) Catch bond interaction between cell-surface sulfatase Sulf1 and glycosaminoglycans. Biophys J 108:1709–1717. https://doi.org/10.1016/j.bpj.2015.02.028
20. Vives RR, Crublet E, Andrieu JP, Gagnon J, Rousselle P, Lortat-Jacob H (2004) A novel strategy for defining critical amino acid residues involved in protein/glycosaminoglycan interactions. J Biol Chem 279:54327–54333
21. Henriet E, Jäger S, Tran C, Bastien P, Michelet J-F, Minondo A-M, Formanek F, Dalko-Csiba M, Lortat-Jacob H, Breton L, Vivès RR (1861) A jasmonic acid derivative improves skin healing and induces changes in proteoglycan expression and glycosaminoglycan structure. Biochim Biophys Acta Gen Subj 2017:2250–2260. https://doi.org/10.1016/j.bbagen.2017.06.006
22. Vives RR, Sadir R, Imberty A, Rencurosi A, Lortat-Jacob H (2002) A kinetics and modeling study of RANTES(9-68) binding to heparin reveals a mechanism of cooperative oligomerization. Biochemistry (Mosc.) 41:14779–14789
23. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815. https://doi.org/10.1006/jmbi.1993.1626
24. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486
25. Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125:1731–1737. https://doi.org/10.1021/ja026939x
26. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ (2016) The HADDOCK2.2 Web server: user-friendly integrative modeling of biomolecular complexes. J Mol Biol 428:720–725. https://doi.org/10.1016/j.jmb.2015.09.014
27. Ai X, Do AT, Kusche-Gullberg M, Lindahl U, Lu K, Emerson CP Jr (2006) Substrate specificity and domain functions of extracellular heparan sulfate 6-*O*-endosulfatases, QSulf1 and QSulf2. J Biol Chem 281:4969–4976
28. Cardin AD, Weintraub HJ (1989) Molecular modeling of protein-glycosaminoglycan interactions. Arteriosclerosis 9:21–32
29. Walhorn V, Möller A-K, Bartz C, Dierks T, Anselmetti D (2018) Exploring the sulfatase 1 catch bond free energy landscape using Jarzynski's equality. Sci Rep 8:16849. https://doi.org/10.1038/s41598-018-35120-0

2. Substrate Specificity

Since our data showed that HSulf2ΔHD could exhibit residual endosulfatase activity, we sought that maybe the removal of HD prevent the recognition of long chains but not small ones. We thus decided to analyze the ability of both HSulf-2 WT and HSulf2ΔHD to desulfate small oligosaccharides.

a. Minimal size for substrate recognition

For this, oligosaccharides of different sizes [IdoA(2S), GlcNS(6S)]$_n$ (where n=1 for disaccharide, n=2 for tetrasaccharide, n=3 for hexasaccharide and n=4 for octasaccharide) (see page 178), were treated with HSulf-2 or with HSulf2ΔHD for 1h, 4h, 7h, 24h, 30h and 48h. The oligosaccharides were then digested with heparinases I, II, III to generate disaccharides that were analyzed by RPIP-HPLC.

Regarding HSulf-2 activity, results showed that, within few hours, all the sulfate groups were removed from the octa, hexa and even tetrasaccharide (Figure 31). No digestion was observed for the disaccharide even after 48h (data not shown), meaning that a tetrasaccharide is the minimal oligosaccharide size required for HSulf-2 activity.

Regarding HSulf2ΔHD, it was able to digest the oligosaccharides. The HD domain is thus not required for the specific recognition of small HS substrate by the enzyme. Interestingly, none of the oligosaccharide digestion did go to completion, even after 48h. HSulf2ΔHD actually removed only one sulfate group from each oligosaccharides (Figure 31). These data thus indicate that HSulf2ΔHD has an “exosulfatase-like” activity and that the HD domain is essential for the processivity of enzyme along the HS chain. Noteworthy, when we highly increased the concentration of the enzyme, the enzyme was able to remove more than one sulfate group of the octasaccharide (data not shown). These observed data showed that the HSulf2ΔHD is not exclusively an exoenzyme, but an enzyme with impaired processivity.

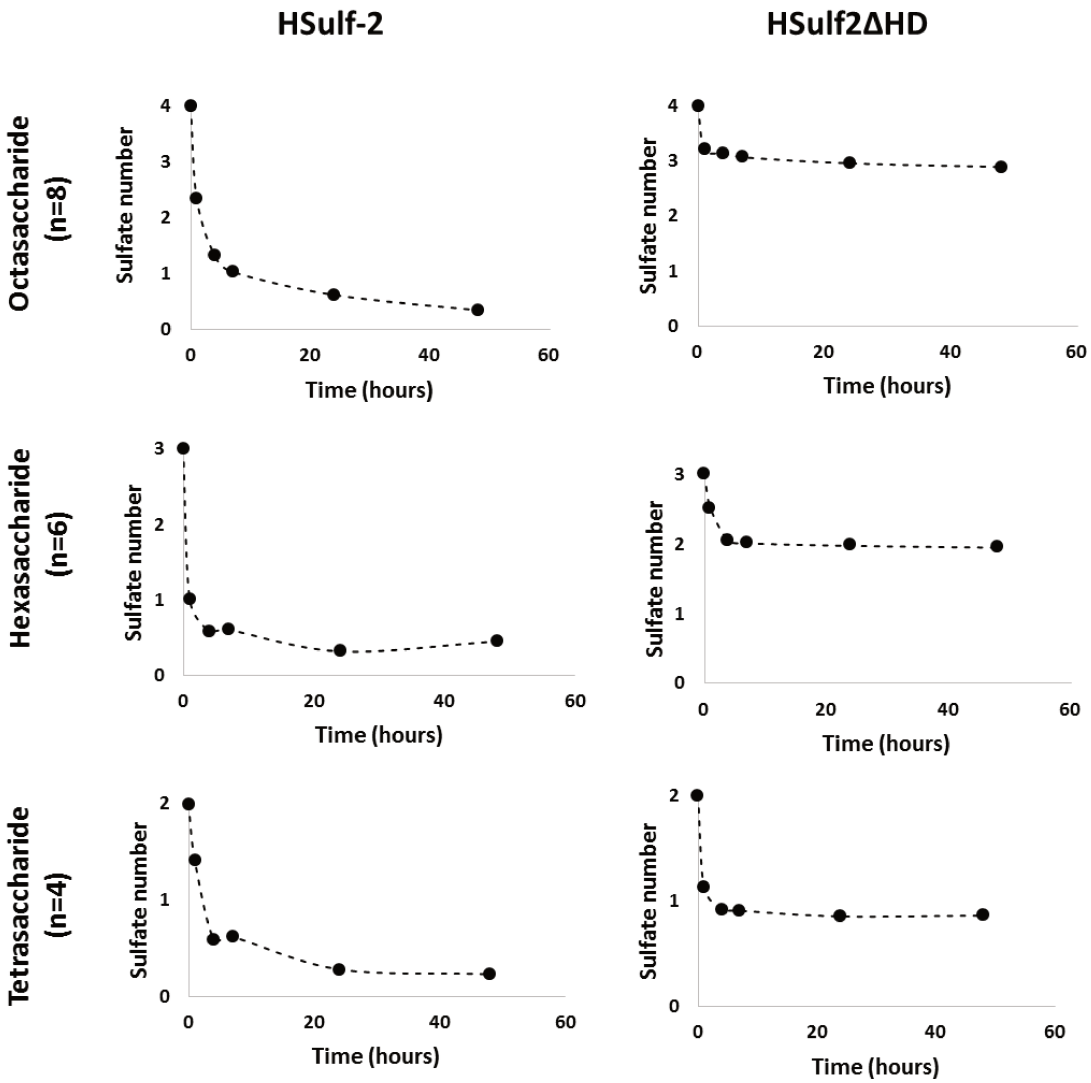


Figure 31: Activity of HSulf-2 and HSulf2ΔHD on oligosaccharides. The evolution of the number of 6-*O*-sulfate groups within time of oligosaccharides [IdoA(2S), GlcNS(6S)]$_n$ digested with HSulf-2 or HSulf2ΔHD. The sulfate number is calculated based on the percentage of disaccharide substrates. For example, an octasaccharide comprises four [IdoA(2S), GlcNS(6S)] disaccharides, each featuring one 6-*O*-S groups. A 25% decrease in disaccharide substrate following digestion thus corresponds to the loss of one 6-*O*-S group.

b. Structural requirements at the oligosaccharide non-reducing end

We next wanted to investigate the importance of the presence of an uronic acid residue at the oligosaccharide non-reducing end for HSulf-2 activity. For this, the [IdoA(2S), GlcNS(6S)]$_3$ hexasaccharide was treated with mercuric acetate in order to generate a pentasaccharide lacking the unsaturated non-reducing end uronic acid. The digestion was confirmed by RPIP-HPLC (Figure 32), as a shift in the peak elution time was observed, corresponding to the loss of the uronic acid charges (one sulfate and one carboxyl group). The resulting pentasaccharide was then treated with HSulf-2 for 48h and analyzed by RPIP-

HPLC. A new peak shift was observed, showing a further reduction of the pentasaccharide. We could thus conclude that HSulf-2 could successfully remove 6-*O*-sulfate groups from a non-reducing end uronic acid deprived pentasaccharide.

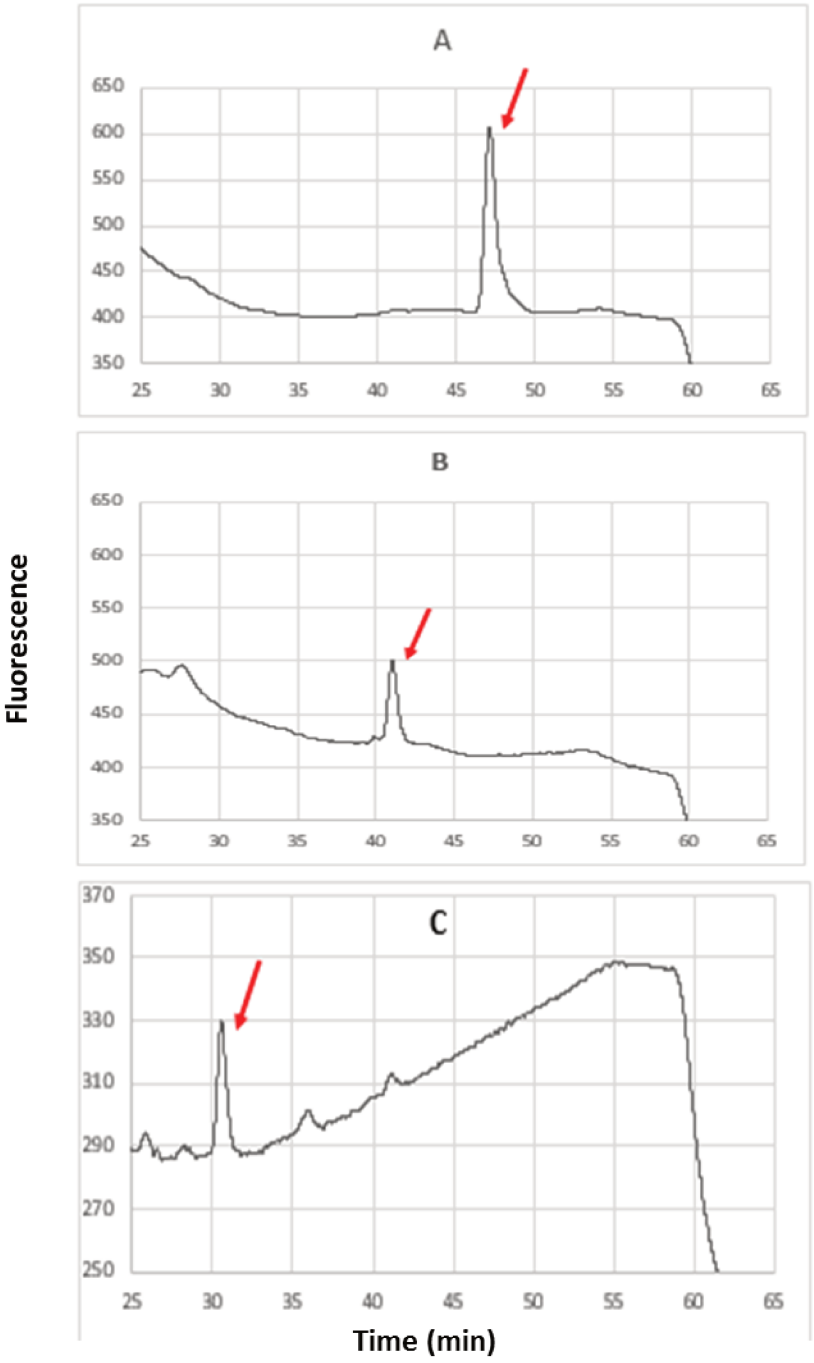


Figure 32: HSulf-2 digestion of pentasaccharide. The elution profile within time of non-treated [IdoA(2S), GlcNS(6S)]$_3$ hexasaccharide (A), mercuric acetate-treated hexasaccharide to generate a pentasaccharide (B), and HSulf-2 treated pentasaccharide (C).

c. Digestion of synthetic oligosaccharides

We next aimed at studying the importance of the 6-*O*-sulfate group position for the enzyme activity. For this, we used a series of synthetic octasaccharides (provided by Christine Le Narvor and David Bonnaffé, Orsay University, France) with similar $[IdoA(2S), GlcNS(6S)]_8$ structure, but with one 6-*O*-S group of the GlcNS missing on each of the disaccharide units of the oligosaccharide (Figure 33, see page 179).

The oligosaccharides were treated with HSulf-2 or HSulf2ΔHD for 1h, 4h, 7h, 24h, 30h and 48 hours and the digestion products were analyzed by RPIP-HPLC as before. Results showed that HSulf2ΔHD could remove a sulfate group from 8.2 and 8.3, but not from 8.1 (Figure 33). Given that the non-reducing end 6-*O*-S of 8.1 is already missing, these results confirm the removal of only the first 6-*O*-S by HSulf2ΔHD.

In contrast, HSulf-2 is able to remove at least one sulfate group from all the octasaccharides, even the 8.1. This confirms the endosulfatase activity of the enzyme. It should be noted that results shown for HSulf-2 are preliminary and were obtained using a poorly active batch of enzyme. We hypothesize that this could explain the low levels of digestion achieved. This experiment should be repeated, to assess whether the HSulf-2 can catalyze extensive 6-*O*-desulfation of the synthetic oligosaccharides. Noteworthy, synthetic oligosaccharides are reduced (see page 179). Consequently, their reducing end disaccharide unit cannot be derivatized by 2-cyanoacetamide and detected following RPIP-HPLC (see Material and Methods from CMLS article, see page 90). For this reason, it is not possible to know if HSulf-2 is able to reach the terminal 6-*O*-sulfate group in the 8.3 octasaccharide, despite the absence of 6-*O*-sulfate in position 3. We expect to address this issue using and additional synthetic octasaccharide (8.4, see Figure 33), lacking the 6-*O*-S at position 2, which will be provided by our collaborators.

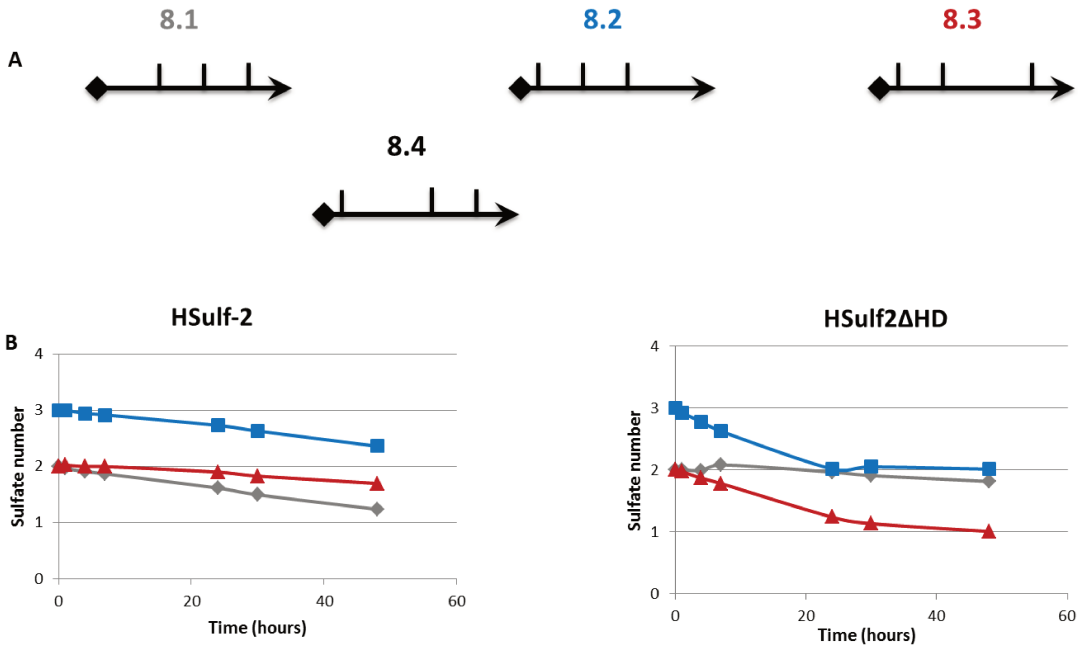


Figure 33: Activity of HSulf-2 and HSulf2ΔHD on synthetic oligosaccharides. Schematic representation of the synthetic octasaccharides [IdoA(2S), GlcNS(6S)]8, lacking one 6-*O*-sulfate group. Each vertical bar corresponds to a 6-*O*-S (A). The evolution of the 6-O-S number within time of synthetic octasaccharides digested with HSulf-2 or HSulf2ΔHD (B).

CHAPTER

IV

In vivo study of HSulf-2

1. Preface

The second functional aspect of this thesis aims at understanding the implication of HSulfs in cancer *in vivo*. Amal Seffouh (former PhD student) previously showed that HSulf-2 exhibited a unique PTM: a CS chain covalently attached to the HD domain of the enzyme. Following this observation, we suggested that this modification could affect enzyme/substrate recognition as well as enzyme diffusion and bioavailability within tissues. One of my objectives was thus to determine the biological role of this chain *in vitro* and *in vivo*.

We first generated a mutant form of HSulf-2 devoid of GAG chain (point substitution of $S_{583}A$ GAG anchorage site). For the *in vitro* study, we compared the activity of the WT and mutant HSulf-2, and their ability to bind to HS. Results showed that the CS chain decreased the binding of HSulf-2 to HS, reducing thus the enzyme endosulfatase activity.

For the *in vivo* study, I was trained at University of Grenoble Alpes, to participate to the *in vivo* experiments, which were performed at the BIG-BCI, CEA, in collaboration with Odile Filhol-Cochet. During this formation, I learned the ethical rules that should be respected in *in vivo* project proposals and during actual animal experiments. I also learned the basics needed to manipulate mice.

To study the consequences of HSulf-2 GAG chain on tumor growth, we first established cancer cell lines expressing stably HSulf-2 WT or a HSulf-2 mutant form that were used in a xenograft model in mice. We chose human breast cancer cell line MDA-MB-231, as these cells do not express HSulfs endogenously (Peterson et al., 2010). Unexpectedly, transfection of these cells with HSulf-2 coding plasmids proved extremely difficult. Classical transfections using transfections reagents such as lipofectamin systematically failed, as cells transiently produced the enzyme, but either died or lost expression after several passages. Use of alternative transfection conditions was also unsuccessful (changing the plasmid, the cell type, the mode of transfection). Alternatively, we attempted direct injection of recombinant enzyme into tumors. Results suggested that HSulf-2 reduced the size of tumors (data not shown). However, we abandoned this approach, as data showed no statistical significance, possibly because amounts of enzyme precisely injected within tumors are difficult to evaluate, and the stability of HSulf-2 in tumors could not be assessed. In addition, these conditions do not mimic the physiological state.

Finally, we changed strategy and eventually succeeded in obtaining stably transfected cells, using a lentivirus system. We next optimized the *in vivo* experimental conditions using the HSulf-2 WT form (the number of injected cells, the medium used, the number and the type of mice). I*n vivo* results showed that the lack of the GAG chain on HSulf-2 increases the capacity of the tumors to grow, vascularize and to metastasize in other tissues. Furthermore, we suggest that the loss of the GAG chain may be governed by a proteolytic process induced in the tumoral environment. The results of this work are presented in the following manuscript, which we expect to submit shortly.

Title

The sweet side of extracellular sulfatases

Authors and Affiliation

Rana El Masri[1$], Amal Seffouh[1$], Caroline Roelants[2], Evelyne Gout[1], Julien Pérard[3], Yoann Crétinon[1], Hugues Lortat-Jacob[1] Odile Filhol-Cochet[2] and Romain R. Vivès*[1]

From [1]Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble, France, [2] Univ. Grenoble Alpes, CEA, Inserm, BIG-BCI, Grenoble, France and [3] Univ. Grenoble Alpes, CNRS, CEA, BIG-LCBM, Grenoble, France.

[*]Correspondence should be addressed to: Romain R. Vivès, IBS, 71 Avenue des Martyrs CS 10090, 38044 GRENOBLE CEDEX 9, France. Phone: (+33) 4.57.42.85.08; Fax: (+33) 4.76.50.19.90; Email: romain.vives@ibs.fr

[$]The authors contributed equally to this work

1. Abstract

Sulfs represent a class of unconventional sulfatases, which differ from all other members of the sulfatase family by their structure, catalytic mechanism and biological function. Through their specific endosulfatase activity, Sulfs provide an original post-synthetic regulatory mechanism of heparan sulfate complex polysaccharides and have been involved in many physiopathological processes, including cancer. However, Sulfs remain very poorly characterized enzymes, with major discrepancies regarding their in vivo functions. Here we show that human HSulf-2 harbors a unique polysaccharide post-translational modification. We identified this glycosylation as a chondroitin/dermatan sulfate (CS/DS) glycosaminoglycan chain, located on the enzyme substrate binding domain. We demonstrated that this CS/DS chain affects enzyme/substrate recognition and tunes HSulf-2 activity in vitro. Finally, we showed that HSulf-2 undergoes proteolytic processing in vivo, leading to loss of the CS/DS chain, and that HSulf-2 lacking the CS/DS chain promoted tumor growth and metastasis in a mouse model. In conclusion, our results highlight HSulf-2 as the first proteoglycan enzyme and its newly-identified GAG chain as a critical non-catalytic modulator of the enzyme activity. We believe this observation should dramatically change the present paradigm on these enzymes and contribute to clarify almost twenty years of conflicting data about their activity.

2. Introduction

Eukaryotic sulfatases have historically been defined as intracellular exoenzymes participating to the metabolism of a large array of sulfated substrates such as steroids, glycolipids, and glycosaminoglycan (GAGs), through hydrolysis of sulfate ester bonds under acidic conditions[1]. However, the field took a dramatic turn two decades ago, with the discovery of the Sulfs[2]. Unlike all other sulfatases, Sulfs were shown to be extracellular endosulfatases that catalyzed controlled 6-*O*-desulfation of cell-surface and extracellular matrix (ECM) heparan sulfate (HS), a polysaccharide with vast protein binding properties and biological functions[3–5]. And unlike all other sulfatases, Sulfs could not be related to a straightforward metabolic function, but rapidly emerged as a novel major regulatory mechanism of HS biological activities, with roles in many physiopathological processes, including embryonic development, tissue homeostasis and cancer[6,7].

Sulfs share a common molecular organization. The furin-processed mature form features a sulfatase-conserved N-terminal catalytic domain (CAT) including the enzyme active site (and notably, the catalytic FGly converted cysteine residue), and a unique highly basic hydrophilic domain (HD), which shares no homology with any known protein and is responsible for high affinity binding to HS substrates. Sulfs display a number of post-translational modifications (PTM)[8]. Furin cleavage[9] and N-glycosylations[10,11] may not be required for the enzyme activity, but play a role in the enzyme attachment at the cell surface, while conversion of C_{88} into a FGly residue is a hallmark of all sulfatases and is essential to the catalytic activity[12]. Studies recently reported that human Sulfs (HSulfs) catalyzed the 6-*O*-desulfation of HS through an original, processive and orientated mechanism[13], and that substrate recognition by the enzyme HD domain involved multiple, highly dynamic, non-conventional interactions [14,15].

However and despite increasing interests, Sulf remain highly elusive enzymes and little is known about their structure, catalytic mechanism and substrate specificity. Our little understanding of these enzymes is well illustrated by the wealth of conflicting data in the literature, reporting major discrepancies between *in vitro* and *in vivo* data, according to the biological system or the enzyme isoforms considered. This is particularly clear in cancer, where both anti-oncogenic and pro-oncogenic effects of the Sulfs have been reported[6,7].

3. Results

HSulf-2 is proteoglycan enzyme

Recently, we reported for the first time high yield expression and purification of Human HSulf-2, which paved the way to progress in the biochemical characterization of Sulfs enzymes [16]. Surprisingly,

size-exclusion chromatography purification step highlighted an unexpectedly high apparent molecular weight (*aMW*) for the enzyme (> ~1000 kDa, for a theoretical molecular weight of 98 kDa), which could not be simply attributed to *N*-glycosylations (Figure 1). Protein aggregation or high order oligomerization were ruled out, as quality control negative staining electron microscopy, which showed purified HSulf-2 as small monodisperse ring-shape particles (data not shown). Noteworthy, we failed to detect the HD-containing C-terminal chain of the enzyme using PAGE analysis, although presence of both chains was ascertained by Edman N-terminal sequencing[16]. In line with this, Seffouh and colleagues reported unusually weak mass spectrometry ionization efficiency of HSulf-2 C-terminal chain[10]. SAXS analysis of the protein yielded Guinier plots in accordance with an extended molecular shape and a MW of ~700 kDa, which supported our previous observations. The *ab initio* determination of the molecular shape of the proteins suggested the presence of 2 entities within the protein: a rather globular domain, which could fit the expected size of HSulf-2 and an elongated most likely flexible moiety (Figure S1). On this basis, we thus speculated that the enzyme could have been purified in complex with HS substrate polysaccharide chains. To test this HSulf-2 was treated with heparinases I II III (to digest HS substrates) or chondroitinase ABC (to digest non-substrate GAGs of CS/DS types) prior to size-chromatography. Results showed no effect of the heparinase treatment (Figure S2), while digestion with chondroitinase ABC dramatically reduced HSulf-2 *aMW* (Figure 1). Attempts to dissociate HSulf-2/CS complexes with urea or high NaCl concentrations showed no effect (Figure S2), thereby suggesting covalent linkage between the polysaccharide and the protein, while chondroitinase treatment enabled immunodetection of the enzyme C-terminal chain by western blot (using anti-HSulf-2 C-terminal antibodies), thus locating the presence of the chain on the HD domain.

GAGs are found covalently bound to specific glycoproteins termed proteoglycans (PGs), through a specific attachment site involving the serine residue of a SG dipeptide motif [17]. Examination of HSulf-2 amino-acid sequence showed two such motifs, located within the enzyme HD domain ($S_{508}G$ and $S_{583}G$). We thus expressed and produced a HSulf-2 mutant form lacking these two motifs (HSulf2ΔSG), which eluted at the same time as chondroitinase treated wild type (WT) HSulf-2, with restored detection of the C-terminal chain by western blot analysis (Figure 1). Single substitutions of the first and second sites (HSulf2ΔSG1 or HSulf2ΔSG2) then confirmed the presence of a CS-type GAG chain on the $S_{583}G$ of the enzyme (Figure S2). Finally, treatment of HSulf-2 transfected HEK 293 cell with xylosides, which are competitive inhibitors of PG glycosylation by priming GAG synthesis[18], significantly inhibited production of the High *aMW* HSulf-2 form (data not shown). Overall, these data provide converging evidence that HSulf-2 features a unique PTM at the level of its HD domain, corresponding to the addition of a CS chain, which thus highlights this enzyme as a member of the large PG family.

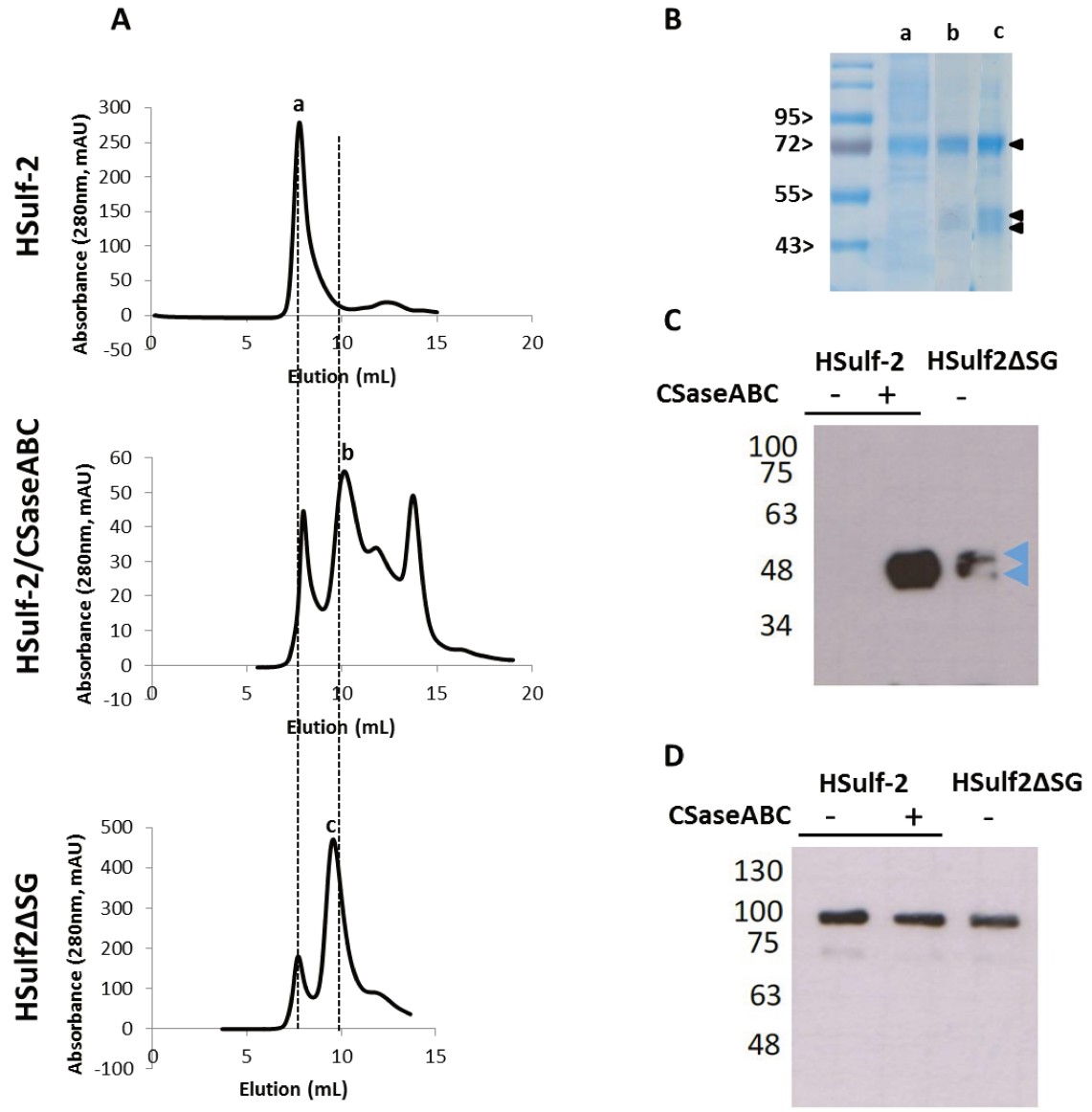


Figure 1 : Expression, purification and characterization of HSulf-2 and HSulf2ΔSG. Size exclusion chromatography profile of HSulf-2, chondroitinase ABC (CSaseABC) pre-treated HSulf-2 and HSulf2ΔSG (A). PAGE analysis of the peaks a, b and c, showing a band of ~ 75 kDa corresponding to the N-terminus of the enzyme, and two bands of ~50 kDa only in peak b and c corresponding the C-terminus (B). Western blot analysis of HSulf-2, CSaseABC pre-treated HSulf-2 and HSulf2ΔSG using anti C-terminal (Abcam) (C) and anti N-terminal (D) antibodies.

HSulf-2 GAG chain modulates enzyme activity *in vitro*

Usually, most biological activities of PGs are attributed to their GAG chains: a PG being a functional enzyme is thus totally unprecedented. To determine the effect of HSulf-2 GAG moiety on its catalytic activity, we compared the enzyme properties of HSulf-2 WT and HSulf2ΔSG forms *in vitro*, using different approaches. We first confirmed that both enzymes exhibited generic arylsulfatase activity

using fluorogenic pseudosubstrate 4-methyl umbelliferyl sulfate (4-MUS) (Figure 2A). We then assessed ability of the enzymes to digest natural substrates using HS analog heparin (HP). Results (Figure 2B) showed enhanced polysaccharide 6-*O*-desulfation with the HSulf2ΔSG form. We thus hypothesized that increased activity could be due to interactions or electrostatic hindrance that could prevent binding of the substrate to the HD domain. To confirm this, we compared the binding of HSulf-2 WT and HSulf2ΔSG to heparin by ELISA (Figure 2C), and to the cell surface HS of human epithelial Wish cells by FACS (Figure 2D). For both approaches, results showed a significant increase in binding for the HSulf-2 form lacking the CS chain. We thus conclude from these data that the *in vitro* activity of HSulf-2 is thus affected by its CS chain that modulates enzyme/substrate binding properties.

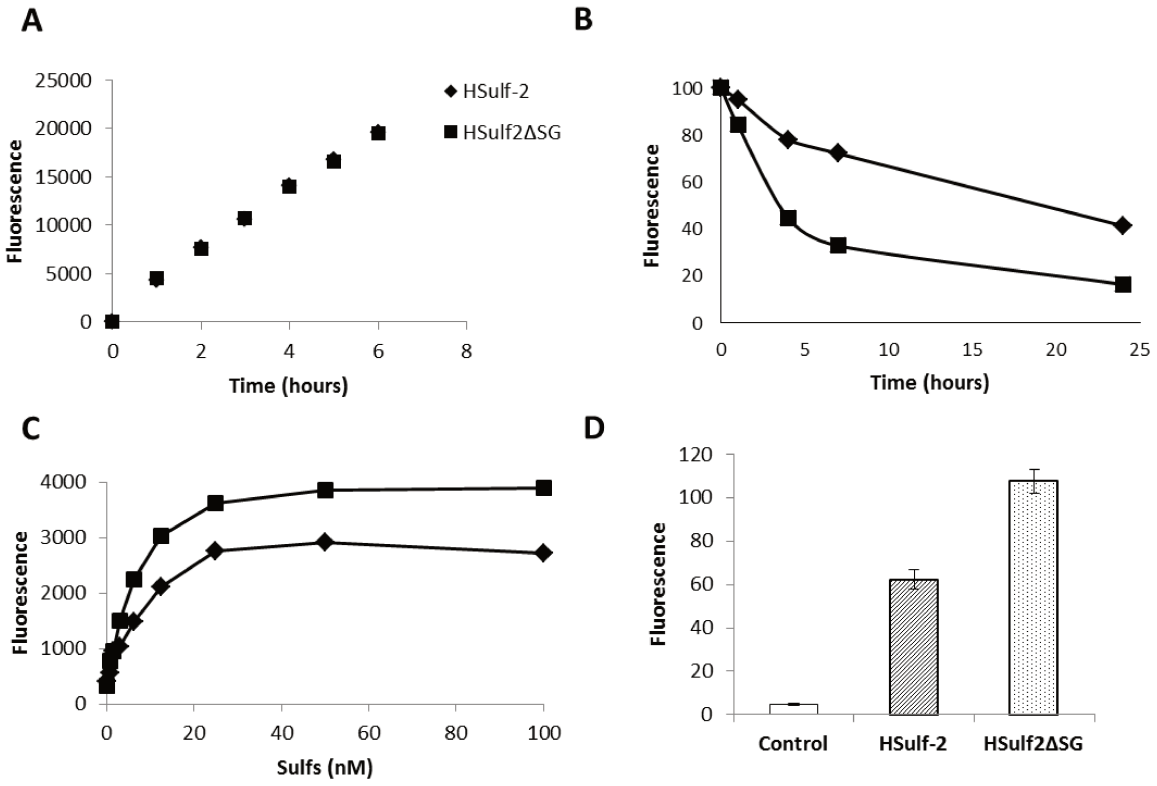


Figure 2 : Comparison of the biological activities of HSulf-2 WT and HSulf2ΔSG mutant. Arylsulfatase assay: fluorescence measurement of 4MUS time course digestion with HSulf-2 WT or HSulf2ΔSG (A). Endosulfatase assay: Time course digestion of Sulf disaccharide in HP digested with HSulf-2 or HSulf2ΔSG (B)[1]. Immuno binding assay of HSulf-2 or HSulf2ΔSG on immobilized HP (C). FACS analysis of HSulf-2 or HSulf2ΔSG binding to the surface of Wish cells using an anti-HSulf-2 antibody. The background fluorescence level of the antibody is shown as a control (D). Data displayed in (C) and (D) are representative of three independent experiments.

[1] The endosulfatase assay was applied based on the calibration of the 4MUS assay. Data need to be repeated in order to have triplicate.

HSulf-2 GAG chain enhances tumor growth and metastasis *in vivo*

To investigate the role of HSulf-2 GAG chain during tumor progression, we overexpressed either HSulf-2 WT or HSulf2ΔSG in MDA-MB-231, a human breast cancer cell line that does not produce any HSulfs endogenously[19]. After selection, HSulf transfected cells exhibited similar expression levels, whereas cells transfected with a plasmid encoding an unrelated protein (DsRed) showed no endogenous expression of the enzyme (Figure S3). Enzyme activity of HSulf-2 produced in MDA-MB-231 was also assessed by treating and analyzing heparin pre-incubated with concentrated condition medium from transfected cells (data not shown). Again, results showed higher endosulfatase activity for HSulf2ΔSG transfected cells. Finally, we confirmed the presence of the GAG chain on MDA-MB-231 HSulf-2 WT (Figure S3).

DsRed, HSulf-2 WT and HSulf2ΔSG transfected MDA-MB-231 cells were then injected into the mammary gland of SCID mice. Tumor volumes were monitored every 2 days and mice were sacrificed when the first tumors reached 1 cm^3 in size (Day 52), in accordance with the European ethical rule on animal experimentation. Tumors along lymph nodes and lungs were then recovered for further analysis. Results showed little effects of HSulf-2 WT expression on tumor size (Figure 3), in contrast with previous work, which reported either anti-oncogenic[19] or pro-oncogenic[20] effects of HSulf-2 WT expression in MDA-MB-231 cells using similar *in vivo* mouse models. Although it should be noted that a major difference between the three studies is the size of xenograft tumors achieved (respectively ~0.04 cm^3 and 3-4 cm^3 in the above mentioned studies), such conflicting data clearly exemplify the complexity of HSulf regulatory functions and possible bias, which could result from the experimental design. In contrast, tumors expressing HSulf2ΔSG mutant increased significantly faster and grew to a larger size, compared to both control and HSulf-2 WT tumors. Noteworthy, HSulf-2 expression levels remained comparable in both HSulf-2 WT and HSulf2ΔSG tumors (Figure S4), thereby indicating that the observed effect on tumor size could be attributed to enhanced *in vivo* activity of the GAG lacking HSulf-2 form. Interestingly, histologyical analysis of tumor sections using eosin/hematoxin coloration showed greater necrotic area in control tumors than in HSulf expressing tumors. As necrosis is a hallmark of hypoxia in growing tumors, we quantified tumor vascularization using α Smooth Muscle Actin (αSMA) immunostaining. Results showed no effect of HSulf-2 WT expression, but significantly increased vascularization in HSulf2ΔSG tumors (Figure 3).

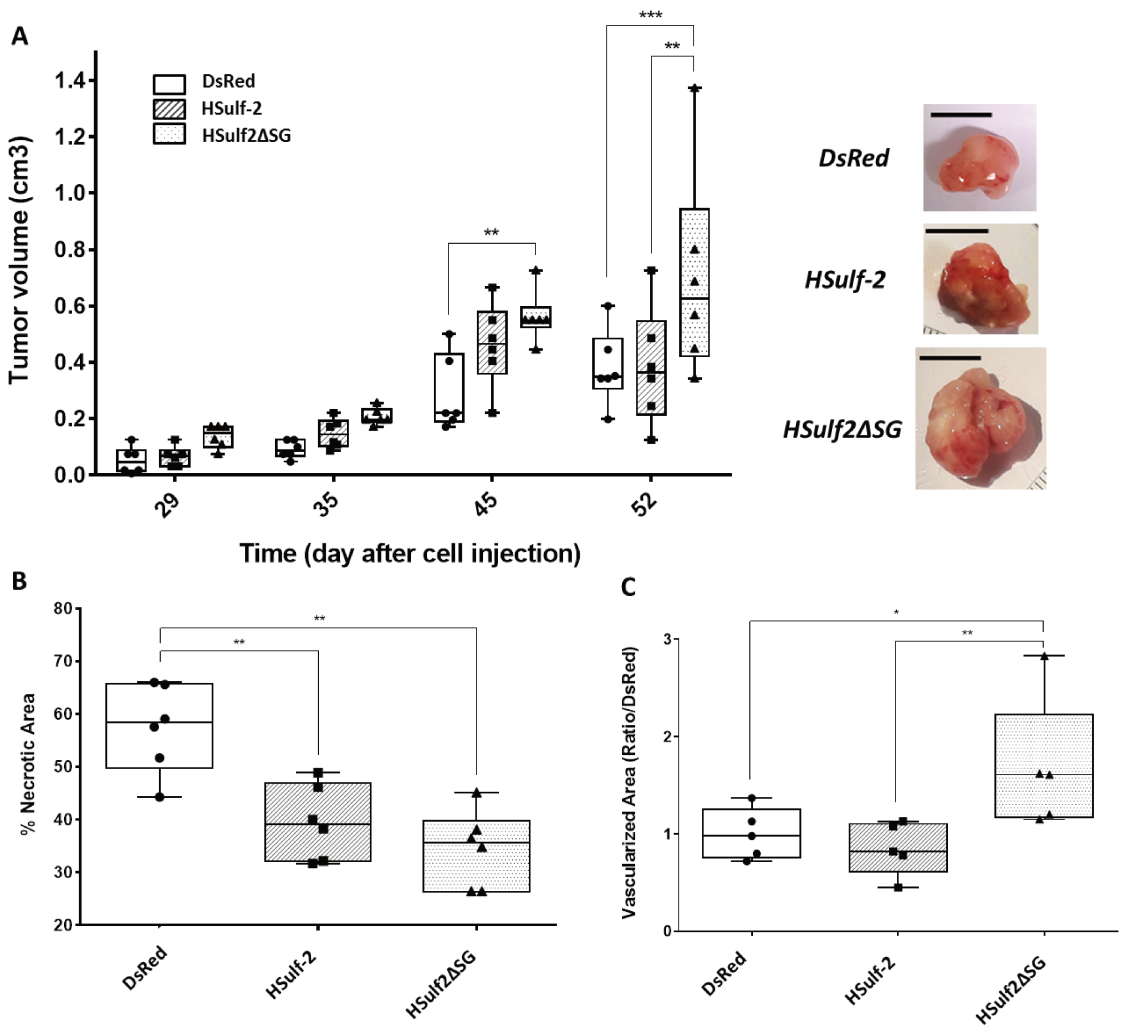


Figure 3: Effects of HSulf-2 WT and HSulf2ΔSG mutant during tumor progression. Time course measurement of the size of tumors induced by MDA-MB-231 cells expressing mock DsRed, HSulf-2 or HSulf2ΔSG (A). Histological analysis of the % of necrotic area using eosin/hematoxin coloration of tumors expressing mock DsRed, HSulf-2 and HSulf2ΔSG. The calculation of necrotic area was performed on three sections from each of the six mice in each group (B). Histological analysis of the vascularized area, using α Smooth Muscle Actin (αSMA) immunostaining of tumors expressing mock DsRed, HSulf-2 and HSulf2ΔSG. The calculation of vascularized area of tumors was performed on five mice in each group (C).

Finally, analysis of tumor invasion showed higher frequency of metastasis in lateral but also control-lateral lymph nodes for HSulfs expressing tumors. Lung, which is a primary target for metastasis in this tumor model, was affected in all conditions. However, metastasis area induced secondary tumors was significantly greater in HSulf2ΔSG expressing tumors (Figure 4). The newly identified CS chain is thus a functionally significant PTM for HSulf-2 biological role in cancer, which attenuates both tumor growth and metastatic invasion *in vivo*.

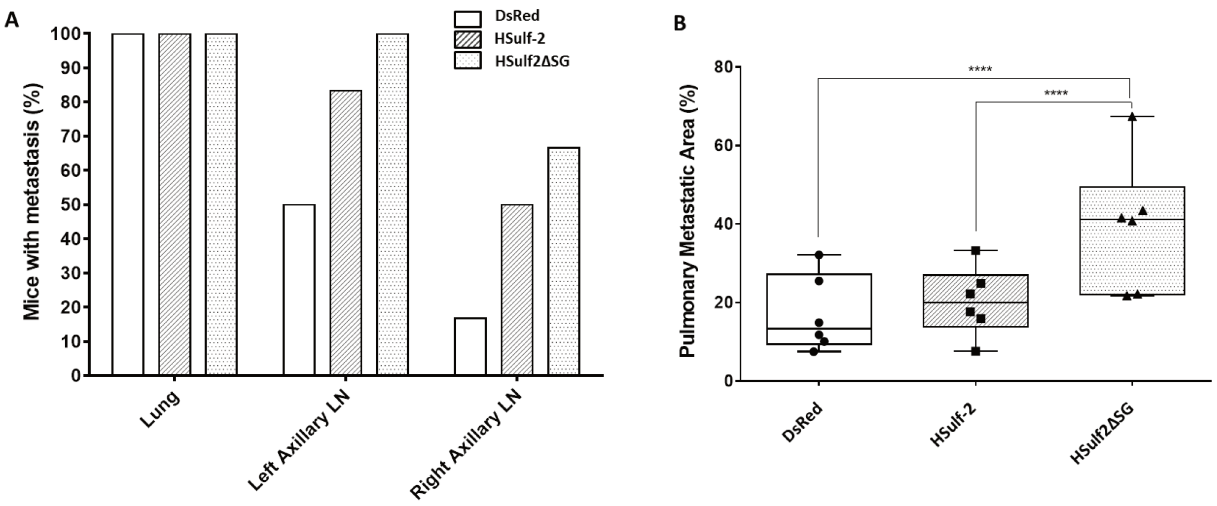


Figure 4: Effects of HSulf-2 WT and HSulf2ΔSG on tumor metastasis. Number of mice with metastasis found in the lung, left axillary and right axillary lymph node (LN) from mock DsRed, HSulf-2 and HSulf2ΔSG expressing tumors (A). Histological analysis of the % of pulmonary metastatic area from mock DsRed, HSulf-2 and HSulf2ΔSG expressing tumors. The calculation of metastatic area percentage was performed on three sections from each of the six mice in each group (B).

4. Discussion

In this study we have identified a novel CS PTM on HSulf-2, and demonstrated its functional relevance using tumor progression as an *in vivo* model. We anticipate that this work will open new and unexpected perspectives in the understanding of the enzyme activities, and will clarify some of the up to now reported discrepancies of the literature.

Through its ability to bind and modulate a wide array of structural and signaling proteins, CS/DS has been involved in most major cellular processes. The presence of such a biologically complex polysaccharide on HSulf-2 may thus have wide functional consequences. Here, we have demonstrated that HSulf-2 CS chain regulated enzyme/substrate recognition, most probably by occupying HS binding sites or eliciting electrostatic hindrance. However, as GAGs bind a wide array of cell-surface and ECM proteins, this CS chain could also be involved in the recruitment of other protein partners, and in the regulation of the enzyme localization, diffusion and bioavailability within tissues.

CS PTM of HSulf-2 may also vary according to tissues and physiopathological status, leading to a variety of enzyme variants, such as forms lacking entirely the GAG chain, or harboring a CS chain with tissue specific structural features. In line with this, we demonstrated the absence of GAG chain on the other human isoform HSulf-1 (data not shown), which could thus account for the functional differences between these two enzymes.

To add to this structural and functional diversity, it is worth noting that HSulf-2 HD domain harboring the CS chain is predicted to be unstructured and prone to controlled proteolysis. In line with this, we have previously reported the formation of HSulf-2 degraded forms during the enzyme purification procedure, which could be inhibited with protease inhibitors[16]. Degradations essentially occurred within the enzyme HD domain and yielded active enzymes lacking the CS chains. Furthermore, western blot analysis of HSulf-2 WT mouse tumors using anti HSulf-2 C-terminal antibody also revealed the presence of similar degradation bands, thus supporting the existence of such controlled processing *in vivo*.

PG controlled proteolysis is a well-documented mechanism termed "shedding", which is induced under certain conditions, and elicits the release of GAG chains/peptide conjugates with specific biological activities [21]. On this basis, our observations may provide the first hints of a more complex regulatory mechanism, in which the enzyme is first produced as a low activity GAG bearing form, which can be subsequently activated upon proteolytic cleavage by extracellular enzymes. This hypothesis opens new perspectives, in which the extracellular environment could play a major role in HSulf-2 activation, especially in an inflammatory context.

Our *in vivo* study provides interesting first evidence of this, as we observed an immunodetection of HD degradation bands for the HSulf-2 WT sample in tumors, whereas intact full length HSulf-2 was detected in conditioned medium from MDA-MB-231 transfected cells[2]. Further investigation will be needed to demonstrate the physiological relevance of such process and fully decipher underlying mechanisms.

In conclusion, this study sheds new light on a complex and highly intriguing enzyme, highlights GAG PTM as a novel non-catalytic regulatory element of its activity, and suggests an original mechanism of protein activation through removal of the GAG chain by controlled proteolysis. But most of all, by identifying such a structurally and functionally relevant feature, which had been overlooked for more than 20 years, we believe that this work strongly encourages to reconsider afresh the importance of PTMs in complex enzymatic systems.

5. Materials and methods

Production and purification of recombinant WT and mutant HSulf-2

[2] Data not shown. A final experiment will be performed to gather all the results in one single figure.

The expression and purification of HSulf-2 and mutants were performed as described previously[16]. Briefly, FreeStyle HEK293F cells (Thermo fisher scientific) were transfected with pcDNA3.1 vector encoding for HSulf-2 cDNA flanked by TEV cleavable SNAP and His tags at N- and C-terminus, respectively. The protein was purified from conditioned medium by cation exchange chromatography on a SP sepharose column (GE healthcare) in 50 mM Tris, 5 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.5, using a 0.1-1 M NaCl gradient, followed by size exclusion chromatography (Superdex200, GE healthcare) in 50 mM Tris, 300 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7. Treatment of HSulf-2 with chondroitinase ABC was achieved by incubating 250 µg of enzyme with 100mU chondroitinase ABC (Sigma) overnight at 4°C. HSulf2ΔSG mutants (ΔSG, ΔSG1, ΔSG2) were generated by site directed mutagenesis (ISBG Robiomol platform) and purified as above.

Western Blot analysis

MDA-MB-231 cells were lysed with RIPA buffer for 2 hours at 4°C and tumors were lysed with RIPA buffer in the presence ceramic spheres (MP Bio) using 3 cycles of 30 seconds at 5000 U in MagNA Lyser. Lysates or purified proteins were then separated by SDS-PAGE, followed by transfer onto PVDF membrane. Proteins were probed with primary mouse anti C-terminal HSulf-2 (Abcam, dil. 1/1000 or R&D systems, dil. 1/500), or primary rabbit anti N-terminal HSulf-2 (Santa Cruz, dil. 1/1000) or primary mouse anti-CS (Sigma, dil. 1/1000) antibodies , followed by incubation with HRP-conjugated anti-rabbit or anti-mouse (Jackson ImmunoResearch Laboratories, dil. 1/10000) secondary antibodies.

ELISA

As reported before[16], microtiter plates were first coated with 10 µg/mL streptavidin in TBS buffer, then incubated with 10 µg/mL biotinylated heparin, and saturated with 2 % BSA. All the incubations were achieved for 1 h at RT, in 50 mM Tris-HCl, 150 mM NaCl, pH 7.5 (TBS) buffer. Next, the recombinant protein was added, then probed with primary rabbit polyclonal anti-HSulf-2 antibody (H2.3, gift from K. Uchimura, dil. 1/1000) followed by fluorescent (A488) - conjugated secondary anti-rabbit antibody (Jackson ImmunoResearch Laboratories, dil. 1/500). All these incubations were performed for 2 h at 4 °C in TBS, 0.05% Tween, and were separated by extensive washes with TBS, 0.05% Tween. Finally, fluorescence of each well was measured (excitation 485 nm, emission 535 nm). Results shown correspond to means +/- SD of three independent experiments.

FACS analysis

Wish cells (1 million for each condition) were washed with PBS, 1 % BSA (the same buffer is used all along the experiment), and incubated with 2 µg of HSulf-2 enzymes (2 h at 4°C). After extensive

washing, cells were incubated with primary antibody (H2.3, gift from K. Uchimura, dil. 1/500., 1 h at 4°C), washed again, then with secondary A488-conjugated antibody (Jackson ImmunoResearch Laboratories dil. 1/500, 1 h at 4°C). FACS analysis of cell fluorescence was performed on a MACSQuant Analyzer (Miltenyi Biotec, excitation 485nm, emission 535nm) by calculating median over 25000 events. Data are represented as means +/- SD of three independent experiments.

Small Angle X-ray Scattering

HSulf-2 WT was filtered extemporaneously before each experiment, using size exclusion chromatography as described above. SAXS data were collected at the European Synchrotron Radiation Facility (Grenoble, France) on the BM29 beamline at BioSAXS. Scattering profiles were measured at several concentrations, from 0.5 to 1.5 mg/mL. Data were processed using standard procedures with the ATSAS v2.8.3 suite of programs[22]. The *ab initio* determination of the molecular shape of the proteins was performed as previously described[23]. Radius of gyration (Rg) and forward intensity at zero angle (I(0)) were determined with the programs PRIMUS[24], using the Guinier approximation at low Q value, in a Q.Rg range < 1.5. Porod volumes and Kratky plot were determined using the Guinier approximation and the PRIMUS programs[24]. The radius of gyration and the pairwise distance distribution function P(r) were calculated by indirect Fourier transform with the program GNOM [25]. The maximum dimension (Dmax) value was adjusted in order that the Rg R value obtained from GNOM agreed with that obtained from Guinier analysis. In order to build *ab initio* models, several independent DAMMIF[26] models were calculated in slow mode with pseudo chain option and merged using the program DAMAVER[24].

***In vitro* enzyme activity assays**

Detailed protocols for arylsulfatase and endosulfatase assays have been described elsewhere[16]. For the arylsulfatase assay, the enzyme (1-3 µg) was incubated with 10 mM 4MUS (Sigma) in 50 mM Tris, 10 mM $MgCl_2$ pH 7.5 for 1-4 hours at 37°C, and the reaction was followed by fluorescence monitoring (excitation 360 nm, emission 465 nm). The endosulfatase assay was achieved by incubating 25 µg of heparin with 3 µg of enzymes in 50 mM Tris 2.5 mM $MgCl_2$, pH 7.5, overnight at 37°C. Disaccharide composition of Sulf-treated heparin was then determined by exhaustive digestion of the polysaccharide (48 hours at 37°C) with a cocktail of heparinase I, II and III (10mU each), followed by RPIP-HPLC analysis[27], using NaCl gradients calibrated with authentic standards (Iduron).

Lentiviral transfection of MDA-MB-231 cells.

HSulf-2 (WT and mutant) DNA were cloned in pLVX vector, a lentiviral vector. This vector was then mixed it with two other viral vectors GAG POL (psPAX2) and ENV VSV-G (pCMV). HEK293T were

transduced with this mix in order to produce pseudoviruses released in the extracellular medium. MDA-MB-231 cells were then infected with those pseudoviruses.

MDA-MB-231 cells were purchased from ATCC and were cultured in Leibovitz's medium supplemented with 10 % fetal bovine serum, 100 U/mL penicillin, 100 µg/mL streptomycin. They were transduced with pLVX-DsRed N1 (Clonetech) as control cells or with pLVX vector encoding either HSulf-2 or HSulf2ΔSG cDNAs, and selected with 2 µg/mL puromycin.

***In vivo* experiments**

In vivo experiment protocols were approved by the institutional guidelines and the European Community for the Use of Experimental Animals. 7-weeks-old female NOD SCID GAMMA/J mice were purchased from Charles River and maintained in the Animal Resources Centre of our department. 1 x 10^6 MDA-MB-231 resuspended in 50 % MatrigelTM (Becton Dickinson) in Leibovitz medium were injected into the fat pad of #4 left mammary gland. Tumor growth was recorded by sequential determination of tumor volume using caliper. Tumor volume was calculated according to the formula $V = ab^2/2$ (a, length; b, width). Mice were euthanized after 52 days through cervical dislocation. Tumors and axillary lymph nodes were collected, weighed and either fixed overnight in 4 % paraformaldehyde (PFA) and embedded in paraffin, or stored at -80°C for western blot analysis. Tissue necrosis was assessed by eosin/hematoxylin coloration. For vascularization analysis, sections (5 µm thick) of formalin-fixed, paraffin embedded tumor tissue samples were dewaxed, rehydrated and subjected to antigen retrieval in citrate buffer (Antigen Unmasking Solution, Vector Laboratories) with heat. Slides were incubated for 10 min in hydrogen peroxide H_2O_2 to block endogenous peroxidases and then 30 min in saturation solution (Histostain, Invitrogen) to block non-specific antibody binding. This was followed by overnight incubation, at 4°C, with primary antibody against αSMA (Ab124964, Abcam). After washing, sections were incubated with a suitable biotinylated secondary antibody (Histostain, Invitrogen) for 10 min. Antigen-antibody complexes were visualized by applying a streptavidin-biotin complex (Histostain, Invitrogen) for 10 min followed by NovaRED substrate (Vector Laboratories). Sections were counterstained with hematoxylin to visualize nucleus. Control sections were incubated with secondary antibody alone. Lungs were inflated using 4% PFA and embedded in paraffin. The metastatic burden was assessed by serial sectioning of the entire lungs, every 200 µm. Hematoxylin and eosin staining was performed on lung and lymph nodes sections (5 µm thick). Images were acquired using AxioScan Z1 (Zeiss) slide scanner and quantified using FiJi software.

Experimental data are shown as median ± standard error mean (SEM). Statistical analyses 2way analysis of variance (2way-ANOVA) with multiple comparisons test (GraphPad Prism 6) were used to

evaluate the significance of differential tumor growth and pulmonary metastases (number and area) between three groups of mice ; Mann-Whitney tests were used to evaluate the differential level of necrosis and vascularization inside tumors. A P value of less than 0.05 was considered to be statistically significant. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$.

Acknowledgments

The authors would like to thank Jean-Pierre Andrieu for Edman degradation N-terminal sequencing, Kenji Uchimura for the Anti-HSulf-2 antibody. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). This work was also supported by the CNRS and the GDR GAG (GDR 3739), the "Investissements d'avenir" program Glyco@Alps (ANR-15-IDEX-02), and by grants from the Agence Nationale de la Recherche (ANR-12-BSV8-0023 and ANR-17-CE11-0040) and Université Grenoble-Alpes (UGA AGIR program). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

Author contributions

REM and AS performed most of the biochemical experiments, with from EG and YC. CR, OFC and REM performed *in vivo* experiments and data processing. JP performed SAXS analysis and modeling. RV, OFC and HLJ interpreted the data and supervised experimental work. RV and REM wrote the manuscript.

References

1. Hanson, S. R., Best, M. D. & Wong, C. H. Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew Chem Int Ed Engl* **43**, 5736–63 (2004).
2. Dhoot, G. K. *et al.* Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. *Science* **293**, 1663–6 (2001).
3. Sarrazin, S., Lamanna, W. C. & Esko, J. D. Heparan sulfate proteoglycans. *Cold Spring Harb Perspect Biol* **3**, (2011).

4. Li, J.-P. & Kusche-Gullberg, M. Heparan Sulfate: Biosynthesis, Structure, and Function. *Int. Rev. Cell Mol. Biol.* **325**, 215–273 (2016).
5. El Masri, R., Seffouh, A., Lortat-Jacob, H. & Vivès, R. R. The 'in and out' of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate. *Glycoconj. J.* **34**, 285–298 (2017).
6. Rosen, S. D. & Lemjabbar-Alaoui, H. Sulf-2: an extracellular modulator of cell signaling and a cancer target candidate. *Expert Opin Ther Targets* **14**, 935–49 (2010).
7. Vives, R. R., Seffouh, A. & Lortat-Jacob, H. Post-Synthetic Regulation of HS Structure: The Yin and Yang of the Sulfs in Cancer. *Front Oncol* **3**, 331 (2014).
8. Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S. & Rosen, S. D. Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J Biol Chem* **277**, 49175–85 (2002).
9. Tang, R. & Rosen, S. D. Functional consequences of the subdomain organization of the sulfs. *J Biol Chem* **284**, 21505–14 (2009).
10. Seffouh, I. *et al.* Mass spectrometry analysis of the human endosulfatase Hsulf-2. *Biochem. Biophys. Rep.* **18**, 100617 (2019).
11. Ambasta, R. K., Ai, X. & Emerson, C. P., Jr. Quail Sulf1 function requires asparagine-linked glycosylation. *J Biol Chem* **282**, 34492–9 (2007).
12. Dierks, T. *et al.* Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme. *Cell* **121**, 541–52 (2005).
13. Seffouh, A. *et al.* HSulf sulfatases catalyze processive and oriented 6-O-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity. *Faseb J* **27**, 2431–9 (2013).
14. Harder, A. *et al.* Catch bond interaction between cell-surface sulfatase Sulf1 and glycosaminoglycans. *Biophys. J.* **108**, 1709–1717 (2015).
15. Walhorn, V., Möller, A.-K., Bartz, C., Dierks, T. & Anselmetti, D. Exploring the Sulfatase 1 Catch Bond Free Energy Landscape using Jarzynski's Equality. *Sci. Rep.* **8**, 16849 (2018).
16. Seffouh, A. *et al.* Expression and purification of recombinant extracellular sulfatase HSulf-2 allows deciphering of enzyme sub-domain coordinated role for the binding and 6-O-desulfation of heparan sulfate. *Cell. Mol. Life Sci. CMLS* **76**, 1807–1819 (2019).
17. Esko, J. D. & Zhang, L. Influence of core protein sequence on glycosaminoglycan assembly. *Curr. Opin. Struct. Biol.* **6**, 663–670 (1996).
18. Chua, J. S. & Kuberan, B. Synthetic Xylosides: Probing the Glycosaminoglycan Biosynthetic Machinery for Biomedical Applications. *Acc. Chem. Res.* **50**, 2693–2705 (2017).
19. Peterson, S. M. *et al.* Human Sulfatase 2 inhibits in vivo tumor growth of MDA-MB-231 human breast cancer xenografts. *BMC Cancer* **10**, 427 (2010).

20. Zhu, C., He, L., Zhou, X., Nie, X. & Gu, Y. Sulfatase 2 promotes breast cancer progression through regulating some tumor-related factors. *Oncol. Rep.* **35**, 1318–1328 (2016).
21. Manon-Jensen, T., Itoh, Y. & Couchman, J. R. Proteoglycans in health and disease: the multiple roles of syndecan shedding. *FEBS J* **277**, 3876–89 (2010).
22. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
23. Pérard, J. *et al.* Structural and functional studies of the metalloregulator Fur identify a promoter-binding mechanism and its role in Francisella tularensis virulence. *Commun. Biol.* **1**, 93 (2018).
24. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
25. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
26. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
27. Henriet, E. *et al.* A jasmonic acid derivative improves skin healing and induces changes in proteoglycan expression and glycosaminoglycan structure. *Biochim. Biophys. Acta* **1861**, 2250–2260 (2017).

Supplementary figures

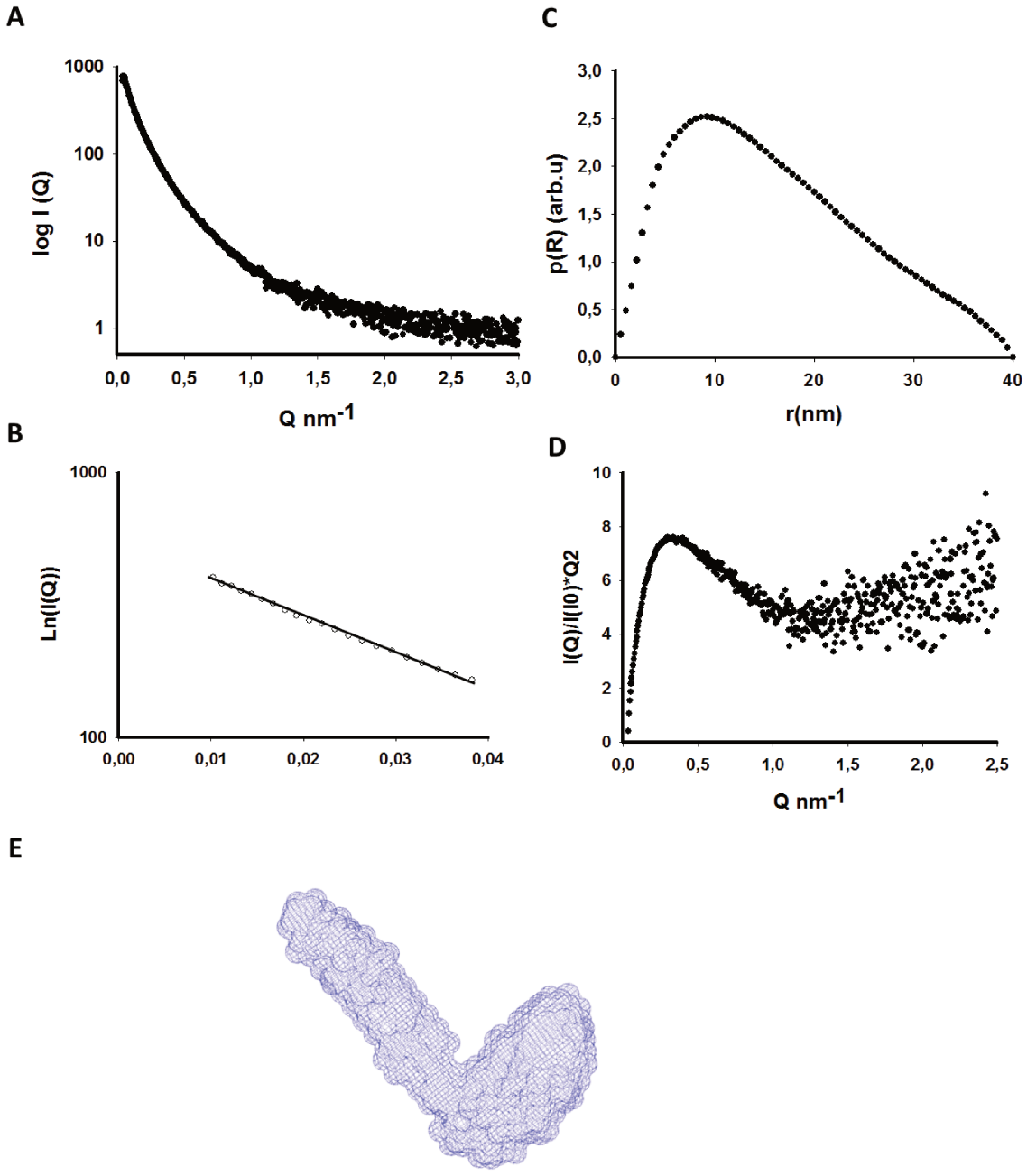


Figure S1. Study of HSulf-2 by Small Angle X-Ray scattering. Scattering curves of experimental data of HSulf-2 in solution at 0.6mg/mL (A). Linear dependence of ln[I(Q)] vs Q2 determined by Guinier plot at 0.6mg/ml with a Rg of 13 nm and I0 of 700 with a porod volume more than 1000. HSulf-2 gives a MW_{exp} : 700 kDa, $MW_{th\ protein}$: 98 kDa (B). Pair distribution function p(R) in arbitrary units (arb.u) *vs. r* (nm) determined by GNOM with a D_{max} of 40nm (C). Kratky plot HSulf-2 shows a partially folded protein with an elongated portion (D). Final *ab initio* model of HSulf-2 generated with 49 individual Dammif model in slow mode and merged with Damaver (NSD < 1) (E).

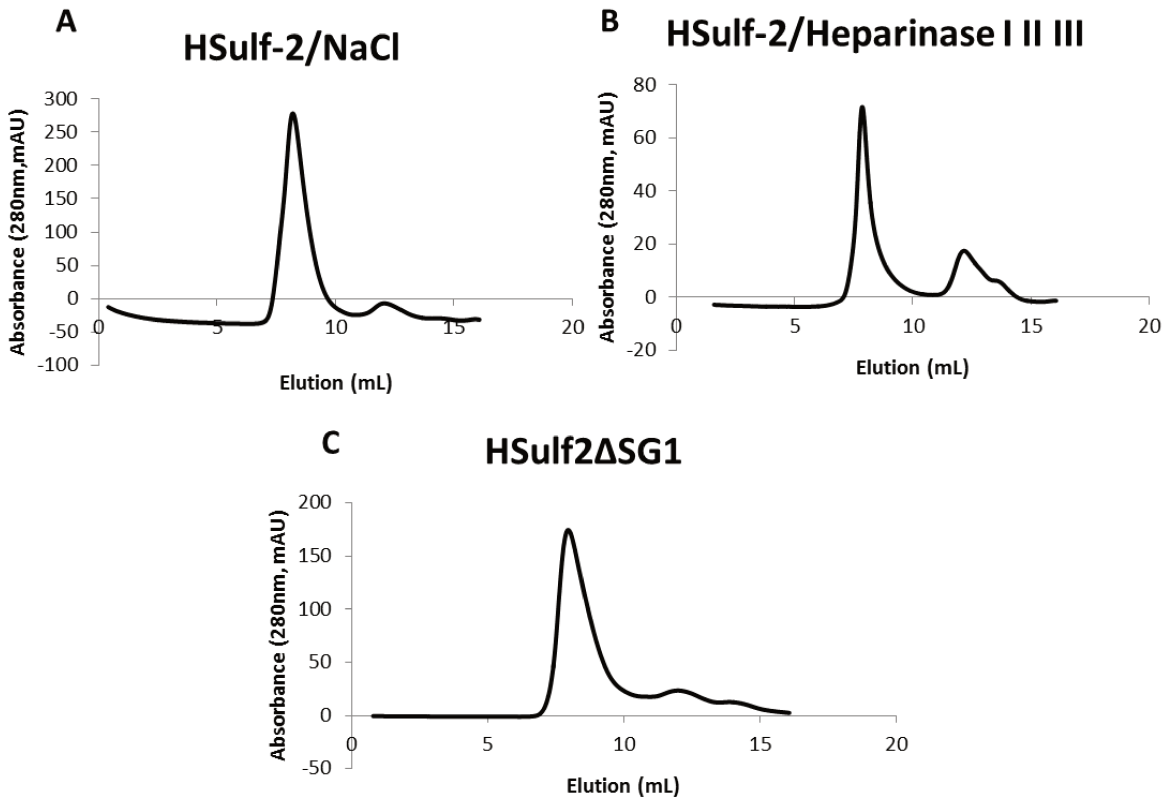


Figure S2 Purification and characterization of HSulf-2. Size exclusion chromatography profile of NaCl pre-treated HSulf-2 (A), heparinase I II III pre-treated HSulf-2 (B) and HSulf2ΔSG1 (C).

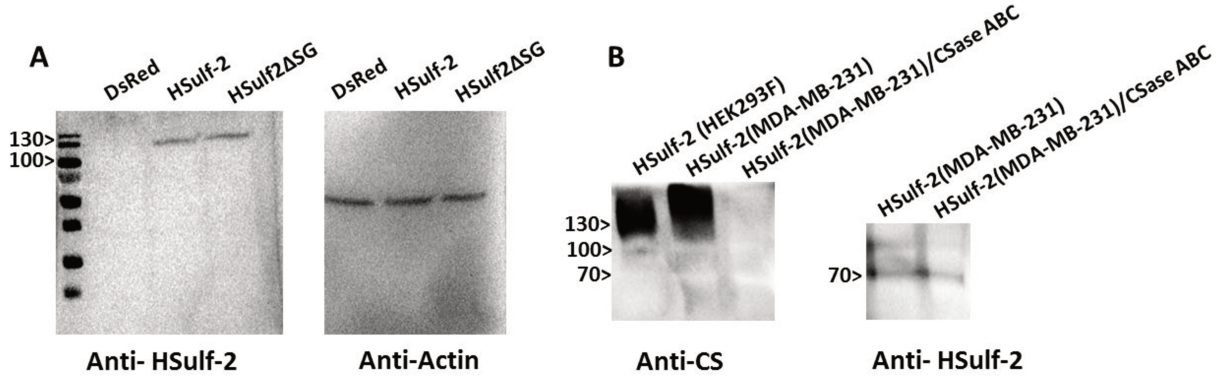


Figure S3: Expression and characterization of Sulfs in transfected MDA-MB-231 cells. Western blot analysis of mock DsRed, HSulf-2 or HSulf2ΔSG transfected MDA-MB-231 cell lysates using anti-HSulf-2 C-terminus antibody (R&D systems) and anti-actin antibody as a control. The 130 kDa band corresponds to the unmaturated whole length protein (A). Western blot analysis of purified HSulf-2 from MDA-MB-231 cells, treated or not with chondroitinase ABC, using anti-CS and anti-HSulf-2 N-terminus (B).

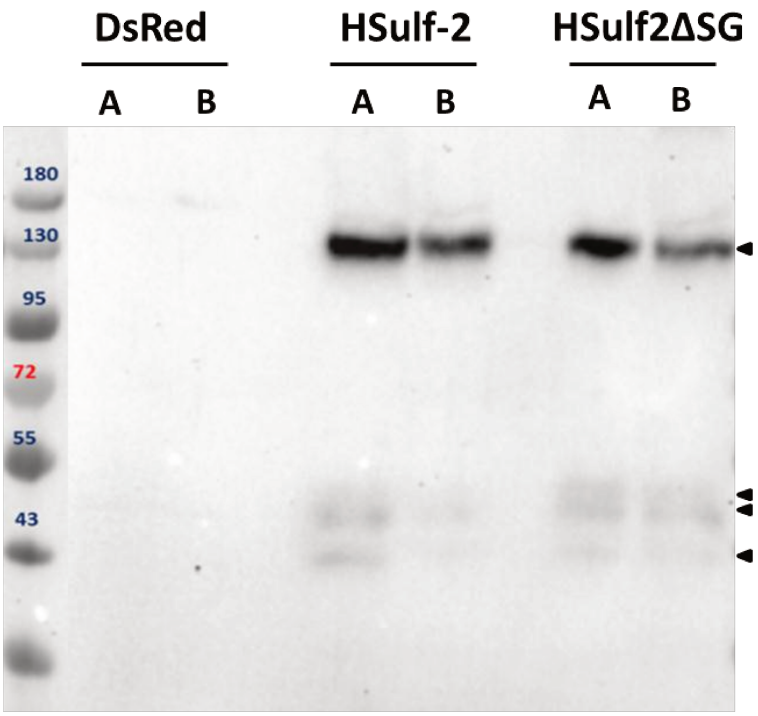


Figure S4 : Detection of Sulfs expression in tumors. Western Blot analysis of mock DsRed, HSulf-2 or HSulf2ΔSG expressing tumor lysates (two tumors A and B per condition) using anti-HSulf-2 C-terminus antibody (R&D systems). The expression levels remained comparable in both HSulf-2 WT and HSulf2ΔSG tumors, based on the 130 kDa band that corresponds to the unmaturated whole length protein. The 43-55 kDa bands correspond to the degradation product of the HSulf-2 HD domain.

CHAPTER

V

Structural study of HSulf-2

1. Assay of crystallization

Before the beginning of my PhD, crystallization assays had been performed on the full length HSulf-2 and were unsuccessful. These failed attempts to obtain protein crystals could be explained by specific structural features of HSulf-2: (i) it is highly N-glycosylated (11 putative N-glycosylation sites); (ii) it comprises the recently identified GAG chain; and (iii) its HD domain is predicted to include a highly disordered and unstructured region of 30 amino-acids residues (based on simulation of disorder score with the iupred software, **Figure** 34). To overcome these difficulties, we tested different strategies, which are discussed below. Unfortunately, none of these approaches led to the formation of diffracting crystals.

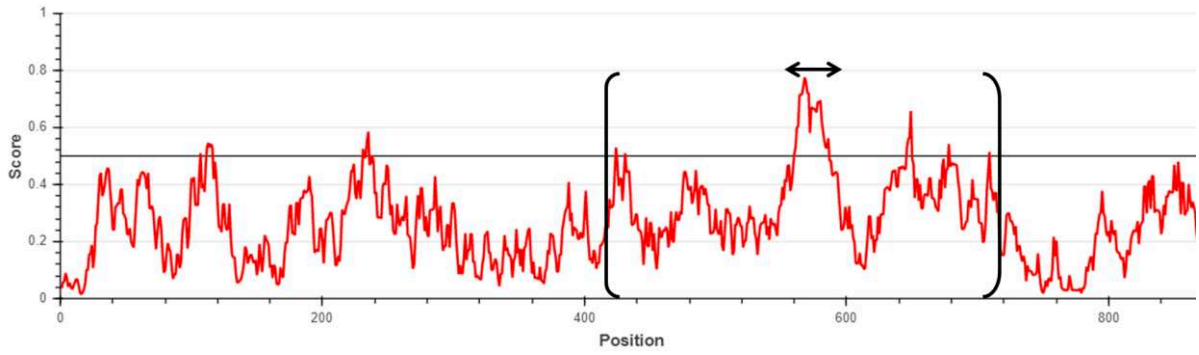


Figure 34: Simulation of disorder score of HSulf-2 by iupred software. The double brackets represent the HD domain of HSulf-2. The arrow corresponds to the disordered regions.

For all crystallization assays, the protein was produced and purified in 50 mM Tris-HCl buffer pH 7.5, 300 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$ as described above (see CMLS article, page 90), then sent to the PSB HTX-lab crystallography platform. Screening conditions for crystallization were typically performed on the enzyme as seated drops, in 1:1 protein/precipitant ratio (100 + 100 nL) in 96-well plates (Qiagen® and Hampton®, 576 conditions tested) and incubated at 20°C, by the nano-volume crystallization robot (Robocrist). The plates were then visualized with the RoboDesign Minstrel III (Rigaku®) over a period of 3 months, taking pictures of the drops regularly, in order to observe and compare their evolution over time.

1. Production of HSulf1ΔHD

We first decided to remove the HD domain from Sulf to avoid any flexibility in the enzyme that could prevent the formation of crystals, given that HD is suggested to be flexible. The rest of the protein is predicted to be structured and is highly homologous to previously crystalized sulfatases, and also among Sulfs isoforms. We started with HSulf-1 isoform,

because HSulf-1 CAT construction design (HSulf1ΔHD) was already established in the literature (Frese et al., 2009). The DNA sequence encoding HD was deleted from HSulf-1 by site directed mutagenesis, as performed previously (Frese et al., 2009). As for HSulf-2 WT, the resulting HSulf1ΔHD cDNA was cloned in a modified pcDNA3.1 vector to add flaking SNAP and His tags at the N- and C-terminus respectively, which could be both removed by TEV treatment. This vector was then used to transfect HEK293F cells, followed by geneticin selection to achieve stable expression. Conditioned medium of HSulf1ΔHD was collected and purified on a nickel column. Fractions were pooled and concentrated to the limit where the protein started to precipitate (~8 mg/ml for HSulf1ΔHD). The imidazole was removed by exchanging the medium using several steps of dilution washes and centricon concentration. The activity of the purified protein was studied. No endosulfatase activity was shown (data not shown), which was expected, since this activity requires the presence of the HD domain on the protein (see CMLS article, page 90). Unexpectedly, no arylsulfatase activity could be detected either (data not shown). However, we previously observed that full length purified HSulf-1 exhibits low aryl-sulfatase activity on 4MUS: this substrate may thus not be the most suitable to assess the activity of our HSulf-1 constructs. In contrast, the purity of the HSulf1ΔHD protein was confirmed by PAGE followed by Coomassie blue staining, and by negative staining electron microscopy (ISBG EM platform - with the help of Daphna Fenel), which showed the presence of homogeneous small size particles of ~7 nm diameter, with no aggregates in the preparation (Figure 35). We also applied Size Exclusion Chromatography coupled to Multi-Angle Laser Light Scattering (SEC-MALLS) study on the protein, in order to confirm to homogeneity of the protein and to determine the experimental MW (ISBG biophysical platform, with the help of Caroline Mas). SEC-MALLS consists on calculating the UV absorbance, the refractive index and the intensity of scattered light, which is directly proportional to the average MW and to the concentration of the sample's components. MW of HSulf1ΔHD was calculated at 98 kDa (in the presence of the SNAP tag). The theoretical MW is 84 kDa, meaning that ~14 kDa may correspond to glycosylations. The sample was sent to the crystallography platform at a concentration of 8 mg/ml, but no crystals could be obtained.

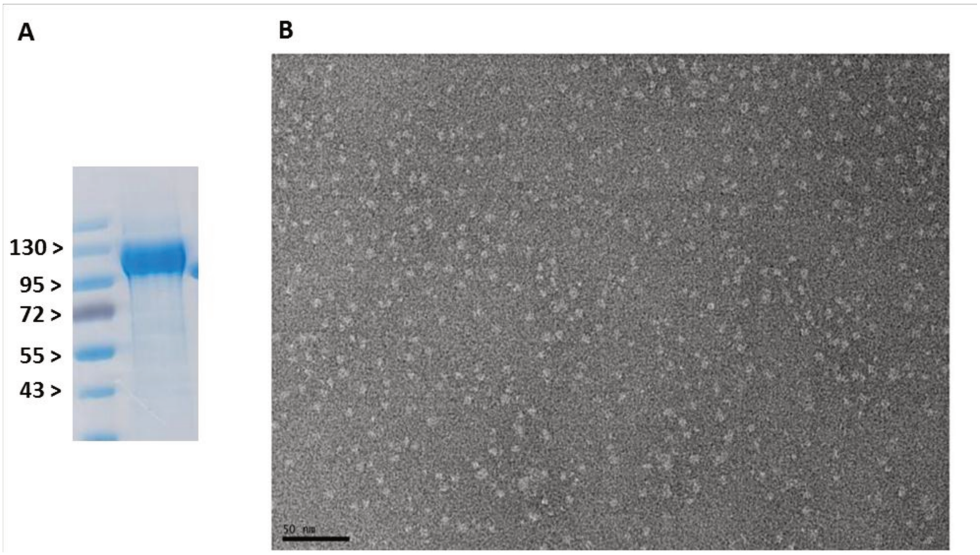


Figure 35: Purification of HSulf1ΔHD. SDS electrophoresis gel of 5µg of purified HSulf1ΔHD stained with blue Coomassie. (A). Negative staining electron microscopy of HSulf1ΔHD (B).

2. Production of HSulf2ΔHD

A similar strategy was applied to achieve preparation of HSulf2ΔHD. The production and the purification of the enzyme were performed as for HSulf1ΔHD. The arylsulfatase activity of the enzyme was confirmed with the 4MUS assay (CMLS article, see page 90). MALDI mass spectrometry performed on HSulf2ΔHD calculated a MW at 107.25 kDa (in the presence of the SNAP tag- ISBG MS platform, with the help of Luca Signor). The theoretical MW is 87.64 kDa, meaning that 19.61kDa may correspond to *N*-glycosylations. The purified enzyme was concentrated to 12 mg/ml and sent to the platform, but did not form crystals.

Synthetic macrocycles such as sulfocalixarenes are known to improve protein assembly and crystallization (Alex et al., 2019). This product is a sulfated molecule that contains phenol groups and could help the stabilization of the protein and thus the crystallization (Figure 36). An additional sample of HSulf2ΔHD protein was sent in the presence of Sulfocalix[4]arene at 1 mM. Again, no crystal could be obtained.

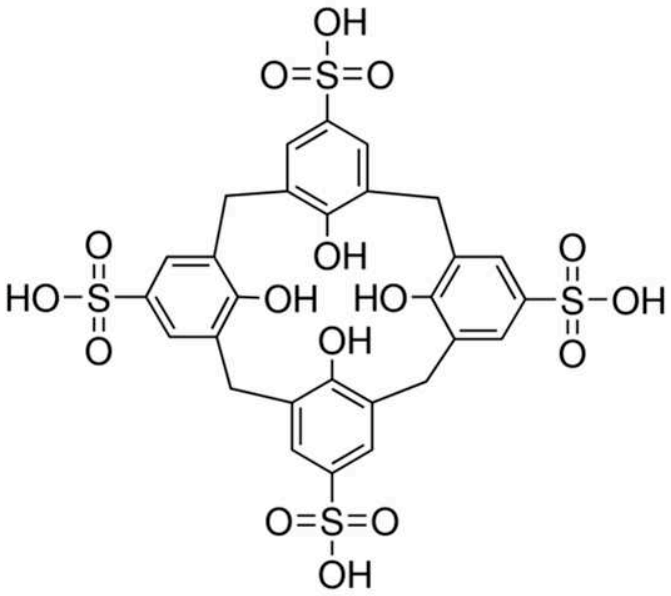


Figure 36: Structure of Sulfocalix[4]arene.

We also hypothesized that absence of crystallization could be due to the presence of the SNAP tag linked to the protein through a small flexible loop. Therefore, in the next assay, conditioned medium of HSulf2ΔHD was collected and purified on a Nickel column, treated with the TEV enzyme to remove both tags and then purified by gel filtration (Superdex200 10/300 GL). Fractions comprising HSulf2ΔHD lacking the SNAP tag were pooled and concentrated to 8 mg/mL. The purity and activity of the protein were confirmed by Coomassie blue stained PAGE and 4MUS assay respectively (data not shown). The sample was sent to the crystallogenesis platform in the presence or not of 10 time excess of sulfocalixarene (1.5 mM). One condition (HSulf2ΔHD without sulfocalixarene) yielded promising primary crystals. However, analysis of the crystal at the ESRF MASSIF beamline showed no diffraction signal.

3. Production of HSulf2ΔHD-S

As glycosylation could increase protein heterogeneity disorder and thus prevent crystallization, we next decided to produce a HSulf2ΔHD form with a reduced level of glycosylations (HSulf2ΔHD-S). To do so, we used HEK-S cells, a mutant of HEK293 cells for N-acetyl-glucosaminyltransferase I (GnTI), which is therefore unable to produce complex saccharide structures beyond that of N-glycan pentasaccharide core. HEK-S cells were transfected with pcDNA3.1/ HSulf2ΔHD, then selected with geniticin. The protein was purified from the conditioned medium of stably expressing cells, using nickel chromatography. Reduction in glycosylation was confirmed by the shift of MW on PAGE,

compared to HSulf2ΔHD expressed in HEK293F (Figure 37). Loss in mass was estimated by MALDI Mass Spectrometry to ~ 6 kDa. We then assessed whether glycan removal could affect the activity of the protein, by performing 4MUS assay on both HSulf2ΔHD and HSulf2ΔHD-S. Results showed that both proteins are active (Figure 37). The enzyme was sent to the platform, but did not form crystals.

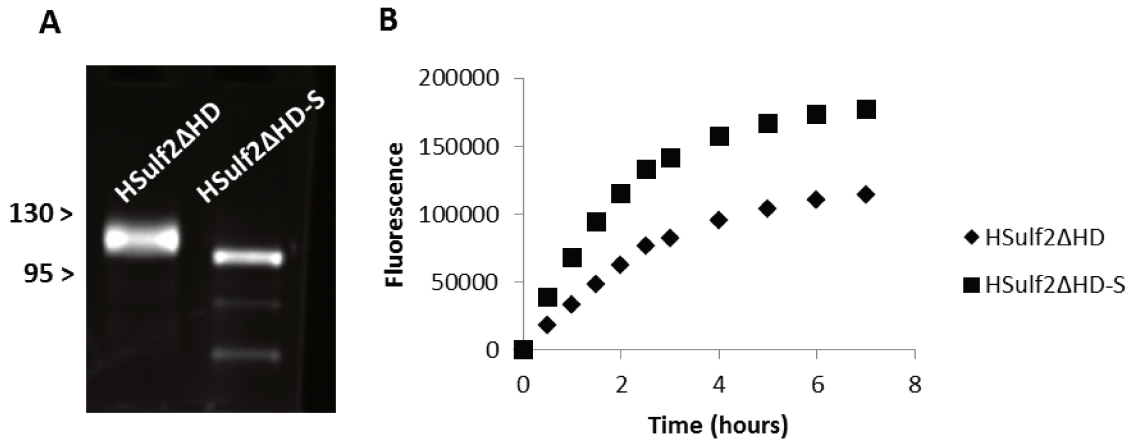


Figure 37 Expression and characterization of HSulf2ΔHD-S. SDS electrophoresis gel revealed by anti-SNAP fluorescent substrate showing HSulf2ΔHD and HSulf2ΔHD-S (A). Fluorescence of the 4-MU within time, resulting from the digestion of 4-MUS with HSulf2ΔHD or HSulf2ΔHD-S (B).

4. Production of Sul-1

In parallel, we also investigated the structure of another ortholog of Sulf: Sul-1 of *Caenorhabditis elegans*, which displays a typical Sulf homologous CAT domain, but a HD domain of reduced size compared to the human isoform. In particular, Sul-1 HD domain seems to lack the unstructured region observed on the human forms, as indicated by sequence analysis using the iupred software (data not shown). This form could thus represent an interesting candidate for crystallization trials. We obtained Sul-1 cDNA from Dr. Tabea Dierker (University of Uppsala, Sweden), and cloned it as previously described. Sul-1 was purified using nickel chromatography, treated by TEV enzyme, followed by size-exclusion chromatography, which yielded satisfactory purity level, as shown by PAGE analysis of the protein preparation (peak a, Figure 38). The peak b was found to be a mix of removed SNAP tag and TEV enzyme (data now shown). The purified protein was concentrated to 3 mg/mL and the activity of the protein was confirmed using the endosulfatase assay, as previously described (data now shown). However, no crystals could be obtained.

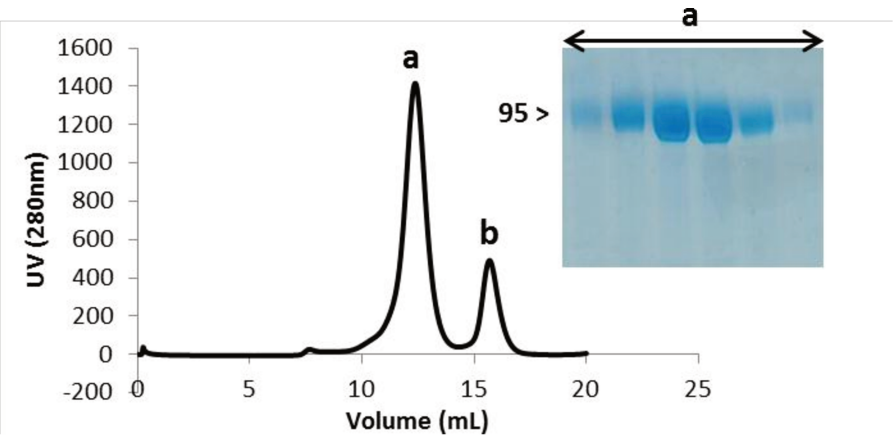
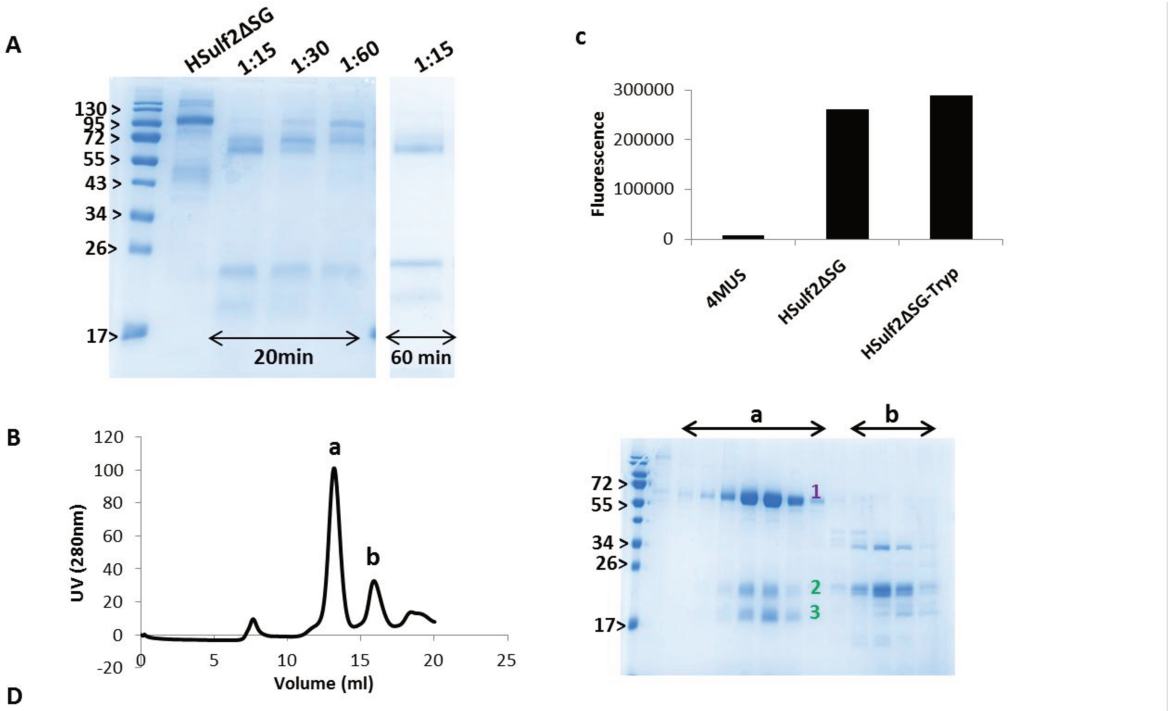


Figure 38: Expression and characterization of Sul-1. Size-exclusion separation profile of Sul-1. Fractions corresponding to the major peak a are indicated by PAGE stained with Coomassie blue.

5. Limited trypsin digestion of HSulf2ΔSG

The strategy of removing the HD domain was not successful to obtain diffracted crystals. The HD may be thus important to maintain the structuration of the whole length protein. For this, we decided to perform controlled digestion of the whole length protein with trypsin, in order to remove the exposed flexible regions that could prevent protein crystallization. We chose to work with the HSulf2ΔSG form that lacks the GAG chain. The protein was incubated with different ratio of trypsin (1:15, 1:30, 1:60) for different times (20min, 60 min), and the result of digestion was analyzed on Coomassie blue stained PAGE (Figure 39). When adding trypsin with 1:60 ratio (1 molecule of trypsin for 60 molecules of Sulf), digestion was observed in both N- and C- terminus. When we increased the concentration of trypsin and the time, we could select the best ratio and incubation time (ratio 1:30, 1 h at 37°C), for which a trypsin-resistant ~ 65 kDa protein core could be obtained. We purified this digestion product (HSulf2ΔSG-Tryp) on size exclusion chromatography (Superdex200 10/300 GL). The profile showed one peak eluted at 13 mL (peak a) and a second one eluted at 16 mL (peak b). These two peaks were analyzed on Coomassie blue stained PAGE. The peak a showed three bands: a major band at ~ 65 kDa, and two minor ones with MW between 17 and 26 kDa (Figure 39). N-terminal sequencing performed directly on these bands showed that the major species corresponded to a protein starting with I_{42}RPN, and that the other two shared the same N-terminal sequence L_{714}QNN, which are the residues located directly after the HD domain. The removed residues upon trypsin digestion are thus the SNAP tag with the first residues of HSulf-2, and the HD domain. Based on the PAGE analysis, we concluded that the peak b corresponded to degraded fragments of the HD domain. We thus hypothesized from these

data that the C-terminal domain is attached to the CAT domain and that HD may be a flexible exposed domain, sensitive to the trypsin digestion. Interestingly, 4MUS assay showed that HSulf2ΔSG-Tryp was active, meaning that the trypsin digestion did not affect the active site of the enzyme (Figure 39). We sent HSulf2ΔSG-Tryp to the crystallography platform at a concentration of 3.6 mg/ml. One crystal could be obtained, but diffraction assay at the Soleil synchrotron yielded a typical diffraction pattern of salt. Nevertheless, the production and the purification of HSulf2ΔSG-Tryp as an isolated protein may be a promising approach for future crystallization assays.



D

RSDKDCEMKRTTLDSPLGKLELSGCEQGLHEIKLLGKGTSAADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFPVPALHHPVFQQ
ESFTRQVLWKLLKVVKFGEVISYQQLAALAGNPAATAAVKTALSGNPVPILIPCHRVVSSSGAVGGYEGGLAVKEWLLAHEGHRLGKPGLGPA
GIGAPGSLGGGSENLYFQSDIFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQVMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSIL
TGKYVHNHNTYTNNENCSSPSWQAQHESRTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTN
MLQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTIL
DIAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEENFLPKYQRVKDLCQRAEYQTACEQLGQK
WQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLG
DAAQPRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNL
REVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHW
QTAPFWTLGPFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPR
TRNMDLGLKDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEGPGENLYFQGPGGGSHHHHHH

Figure 39: Expression and characterization of Trypsin limited digested HSulf2ΔSG. (HSulf2ΔSG-Tryp). SDS electrophoresis gel of non-treated HSulf2ΔSG or treated HSulf2ΔSG with trypsin using 1:60, 1:30 and 1:15 ratio for 20 min or 60 min (A). Size-exclusion separation profile of HSulf2ΔSG-Tryp. Fractions corresponding to the peak a (HSulf2ΔSG-Tryp) and peak b (contaminants or degradations) are analyzed on PAGE stained by Coomassie blue (B). Processing of 4-MUS after a 4h digestion with HSulf2ΔSG or HSulf2ΔSG-Tryp (C). Protein sequence of HSulf-2. SNAP tag, TEV site and HD domain are represented in orange, pink and grey respectively. The three bands showed on the SDS gel in (B) are N-terminal sequenced and are thus indicated in purple and green (D).

2. NMR study of the HD domain

Structure prediction based on the amino-acid sequence analysis of HSulf-2 HD domain suggests that this region may be partially disordered and flexible. As this precludes structural analysis by X-ray crystallography, we proposed to study this domain separately, using NMR. Major challenges of such an approach are that this technique requires large amounts of protein and necessitates isotopic labeling. For these reasons, we chose to produce the HD domain using either a cell free system or a prokaryote system. The first objective was thus to develop the expression of HD on these systems.

I. Optimization of HD expression and purification

1. Cell free system

Cell free protein expression is a fast way to produce protein because it does not require gene transfection, cell culture or extensive protein purification. DNA sequence encoding the HD domain was first cloned in pivex2.4d vector (ISBG Robiomol platform). This vector is designed to add a His tag at the N-terminus of the protein and is optimized for expression using the cell free system. This vector was then used for cell free protein production, with the help of Lionel Imbert (NMR, IBS). This consisted in mixing the vector with the transcriptional and translational machinery extracted from *E. coli* bacteria, which mainly includes RNA polymerase, nucleotides, ribosomes, tRNA, amino-acids and enzymatic cofactors. The reaction occurred overnight at room temperature, and then a centrifugation was necessary to separate the soluble fraction of the supernatant (SN) from the unsoluble fraction of the pellet (P). Both fractions were analyzed on Western blot using anti His antibody (Figure 40). Results showed that the majority of the expressed protein (37 kDa) was found in the pellet. Another band at 25 kDa was revealed by the antibody and was suggested to be a degraded form of HD (degradation at the C-terminal end). However, protein yields in the soluble fraction were not sufficient to proceed with NMR study.

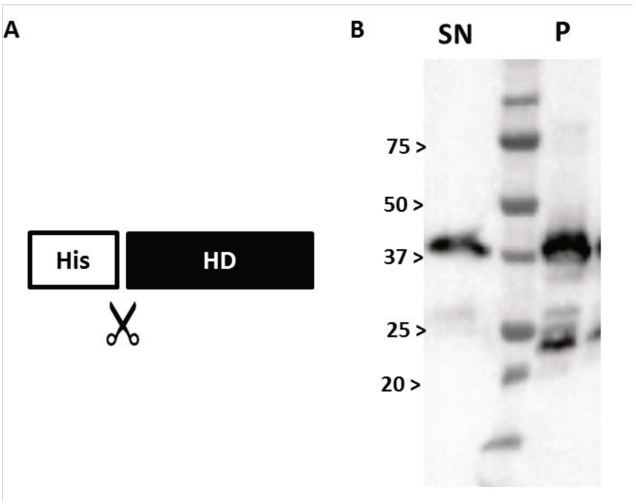


Figure 40: Production of HD-HSulf2 in cell free. Representation of HD-HSulf2 construct. The scissor represents the TEV cleavage site (A). Western Blot analysis of the supernatant (SN) and the pellet (P) fractions resulting from the production of HD-HSulf2 in cell free, using anti-His antibody. The band at 37 kDa corresponds to the HD-HSulf2.

2. Prokaryotic system

We then switched to a prokaryote expression system. For this, we cloned the DNA sequence encoding the HD domain into the pET28a prokaryotic expression vector. This vector was first used to transform different strains of bacteria (BL21, Ril, Rosetta and SHuffle). BL21 is the classical *E. coli* strain suitable for protein expression. Rosetta and Ril are derivatives of BL21 designed to enhance the expression of eukaryotic proteins, given that they supply tRNAs for the codons rarely used in *E. coli*. SHuffle are designed to form protein containing disulfide bonds, given that they express disulfide bond isomerase. The solubility of the produced HD may differ depending on the strain. After 3 hours of induction of protein expression with 1 mM IPTG at 37°C, the bacteria were sonicated and the sonication product was centrifuged to separate soluble and unsoluble forms (inclusion bodies). PAGE analysis of the two fractions showed the presence of the HD in the unsoluble fraction for all the strains (Figure 41). To confirm that HD was actually present within the inclusion bodies and not bound to their surface, the pellet was washed several times with 2 M NaCl and 2 % triton. Analysis of the resulting fractions by PAGE showed that the protein was exclusively found in the pellet (data not shown).

In view of these results, we tested different approaches to improve the expression of HD as a soluble form, which are described below.

a. Growth conditions change

It has been suggested that inclusion bodies are formed when the protein of interest is being produced faster than it can fold into the native structure. For this, we first tried to change growth conditions by using different concentration of IPTG (0.5 mM or 1 mM) and different time and temperature of induction (3 hours at 37°C or overnight at 20°C), as these parameters are known to affect the solubility of the induced protein. However, the protein was still expressed in the inclusion bodies (data not shown).

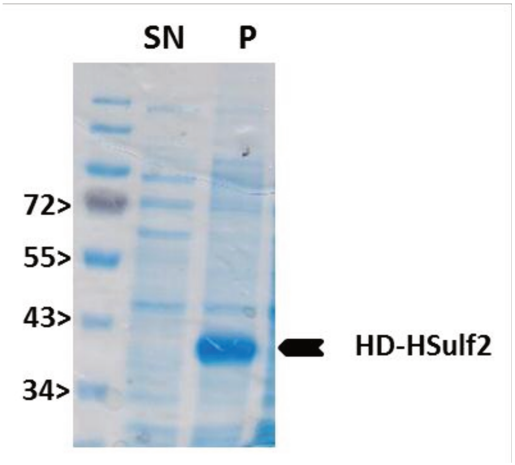


Figure 41: Production of HD-HSulf2 in Ril bacteria. Coomassie blue stained SDS gel electrophoresis analysis of supernatant (SN) and pellet (P) resulting from the sonication of HD-HSulf2 transformed Ril bacteria.

b. Production of HD with a fusion protein

We then decided to change the expression vector and cloned the HD encoding DNA in the pETM-40 vector (provided by Pascal Fender, IBS). This vector enables expression of protein in fusion at the N-terminus with a Maltose Binding Protein (MBP) separated with a TEV cleavage site, which may increase the chance of obtaining a soluble protein (Figure 42). BL21 bacteria were transformed with this vector and expression was induced with 0.5 mM IPTG. Bacteria were then sonicated and the SN and P fractions were analyzed on PAGE. Results showed that the fusion protein (HD + MBP) was produced in important yields, with ~ 50% of the protein found in the soluble fraction. Removal of the MBP protein upon TEV digestion (+TEV) resulted in the formation of a precipitation (Precip). PAGE analysis of the Precip showed that, unfortunately, almost all the HD domain precipated (Figure 42). To avoid this, we tried to perform the TEV digestion in presence of additives, such as Heparin, glycerol or sulfocalixarene to help stabilizing the HD and keeping it soluble. However, none of these conditions worked out (data not shown).

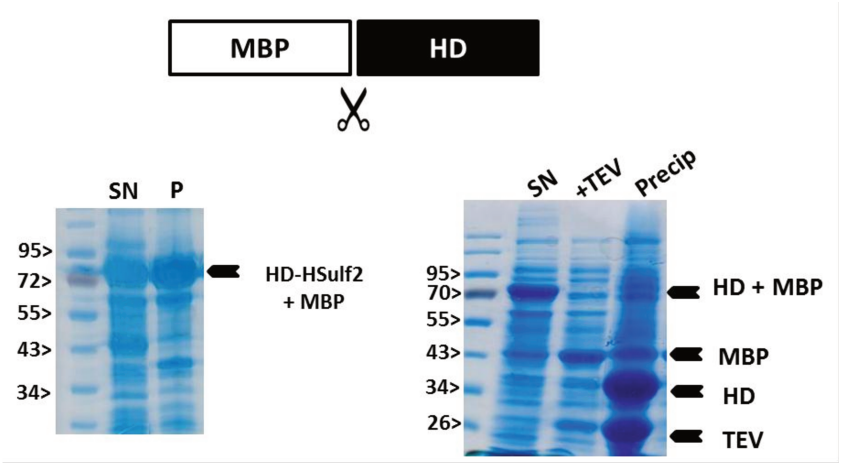


Figure 42: Production of HD-HSulf2 with Maltose Binding Protein (MBP) fusion protein. Representation of HD-HSulf2+MBP construct. The scissor represents the TEV cleavage site (A). SDS gel electrophoresis analysis of supernatant (SN) and pellet (P) resulting from the sonication of HD-HSulf2+MBP transformed BL21 bacteria (B). SDS gel electrophoresis analysis of SN, TEV treated SN (+TEV) and the precipitation resulting from the TEV treatment of SN (Precip) (C).

c. Production of HD in the periplasm

Next, we tried to clone the HD DNA sequence in the pET26b vector (provided by Isabel Ayala, IBS), which adds a signal peptide to the N-terminus of the HD, orientating thus the secretion of proteins towards the periplasm. Unlike the reducing environment of the cytoplasm, the periplasm is an oxidizing environment. It contains disulfide oxidoreductases and isomerases allowing the formation of disulfide bonds, enabling thus the accumulation of properly folded and soluble proteins. As before, BL21 were transformed with pET26b/HD, induced with IPTG, sonicated, and the result of sonication was analyzed on PAGE. Results showed the HD was also found in the inclusion bodies (data not shown).

d. Flash refolding assay

Finally, we decided to use flash refolding techniques in order to solubilize and refold the precipitated proteins from the inclusion bodies. This consisted of solubilizing the inclusion bodies, containing the protein of interest, using a denaturant reagent like 8 M urea or 6 M guanidinium chloride, and then diluting these reagents progressively with an optimized buffer, in which the protein should fold and acquire a functional conformation. For this, we used the initial construction pET28a/HD to transform Ril bacteria. The inclusion bodies containing the HD were solubilized with 8 M urea and the protein was purified in denaturing conditions by nickel chromatography. Next, different techniques of refolding were applied, either directly on the column or using drop by drop dilution in a urea free buffer. Several trials were

performed and ended up with the precipitation of HD after the dilution (data not shown). Noteworthy, the buffer used in these refolding assays was the buffer optimized for stability of the full length HSulf-2 (50 mM Tris, 300 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.5). We thus hypothesized that this buffer could not be suitable for refolding of the isolated HD domain.

In order to determine optimal buffer conditions for HD refolding (and thanks to the advice of Silvia Achili, former PhD student at IBS), we designed a strategy based on a previously published work (Biter et al., 2016). Upon expression, the pellet recovered from centrifugation of the bacteria lysate was solubilized in 6 M guanidinium chloride, 40 mM glutathione, 100 mM Tris, pH 7.5. The protein was purified from this solubilized denaturing fraction by nickel chromatography and concentrated to 2 mg/mL. Flash refolding was then applied in small volume scales by diluting 10 times the denaturated protein with 20 different conditions of buffer, pH and ionic strength. The unfolding transition midpoint Tm was measured (using Differential Scanning Fluorimeter technique, with the help of Michel Thépaut, IBS) and we chose the buffer yielding the highest Tm, which is 52.2°C. The selected optimal buffer was 100 mM HEPES, 150 mM NaCl, 25 mM $CaCl_2$, pH 7 (Figure 43). This buffer was thus applied to a larger batch of denatured protein, which remained soluble after a 10 times dilution.

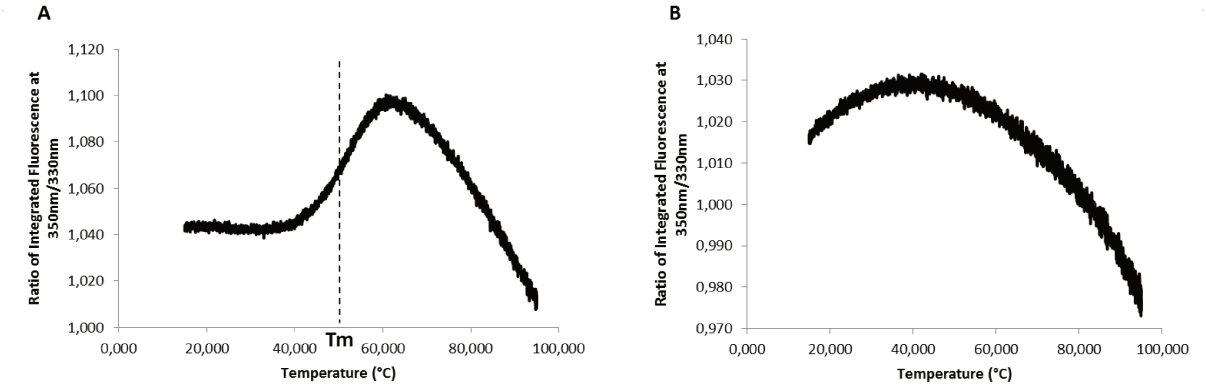


Figure 43: Comparison of the unfolding transition midpoint Tm of HD-HSulf2 in two buffers. 100 mM HEPES, 150 mM NaCl, 25 mM $CaCl_2$, pH 7 (A) or 100 mM citric acid, 150 mM NaCl, pH 4 (B).

To remove completely the guanidinium chloride (600 mM remaining after the 10 times dilution), the protein was next injected onto the size-exclusion column using the selected buffer as running buffer. We pooled the fractions corresponding to the HD peak and

concentrated them but unfortunately, it precipitated upon the concentration (data not shown). The HD was therefore not stable in that buffer.

To solve this problem, we tried to dilute further the 600 mM guanidinium chloride containing protein preparation by dialysis to a final concentration of 6 mM guanidinium chloride before injection onto the size-exclusion column, but this led to the same results (data not show). We next tried to change the ionic strength of the purification buffer by varying the concentrations of NaCl and $CaCl_2$. This last optimization attempts enabled us to recover soluble protein, using a final buffer of 50 mM HEPES, 100 mM NaCl, pH 7. Based on the elution time (9 mL) of the size exclusion chromatography (Superdex75 10/300 GL), we speculated that the protein does not correspond to a monomeric form (Expected MW = 37 kDa), but to a dimeric or multimeric one (Figure 44). Analysis of the protein by PAGE showed a major band of the expected MW, and a minor band that may represent a multimeric state of the protein, given that it was visualized by western blot, using an anti-His antibody and analyzed by N-terminal sequencing indicating HD-HSulf2 sequence (data not shown). The levels of purification using the final optimized protocol vary a lot among purifications, but we can estimate an average of 3 mg/L of culture.

II. Structural analysis of HD

The protein was then analyzed by ^{1}H 1D NMR in order to determine the state of folding of the protein (with the help of Yoan Monneau, former post doc at SAGAG, IBS). This consists of looking at the presence of amide and methyl proton resonances outside the random coil regions. We applied a fast acquisition of few minutes of the HD at 25°C, using a concentration of 150 µM. 1D proton spectrum acquired showed that the amide proton resonances are all in the range 8 – 8.5 ppm, as well as the methyl proton resonances in 0.5 – 1 ppm (Figure 44). This strongly suggests that the HD protein is essentially dynamic or unfolded.

To get further information, we analyzed the HD in ^{1}H ^{15}N 2D NMR (with the help of Yoan Monneau). To do so, the HD was produced in M9 medium using ^{15}N labeling and purified as before. In general, the change of the culture medium from LB to M9 may affect the solubility of the produced protein and the yields of purification. However, in our case, the protein was still unsoluble, and the yields remained the same.

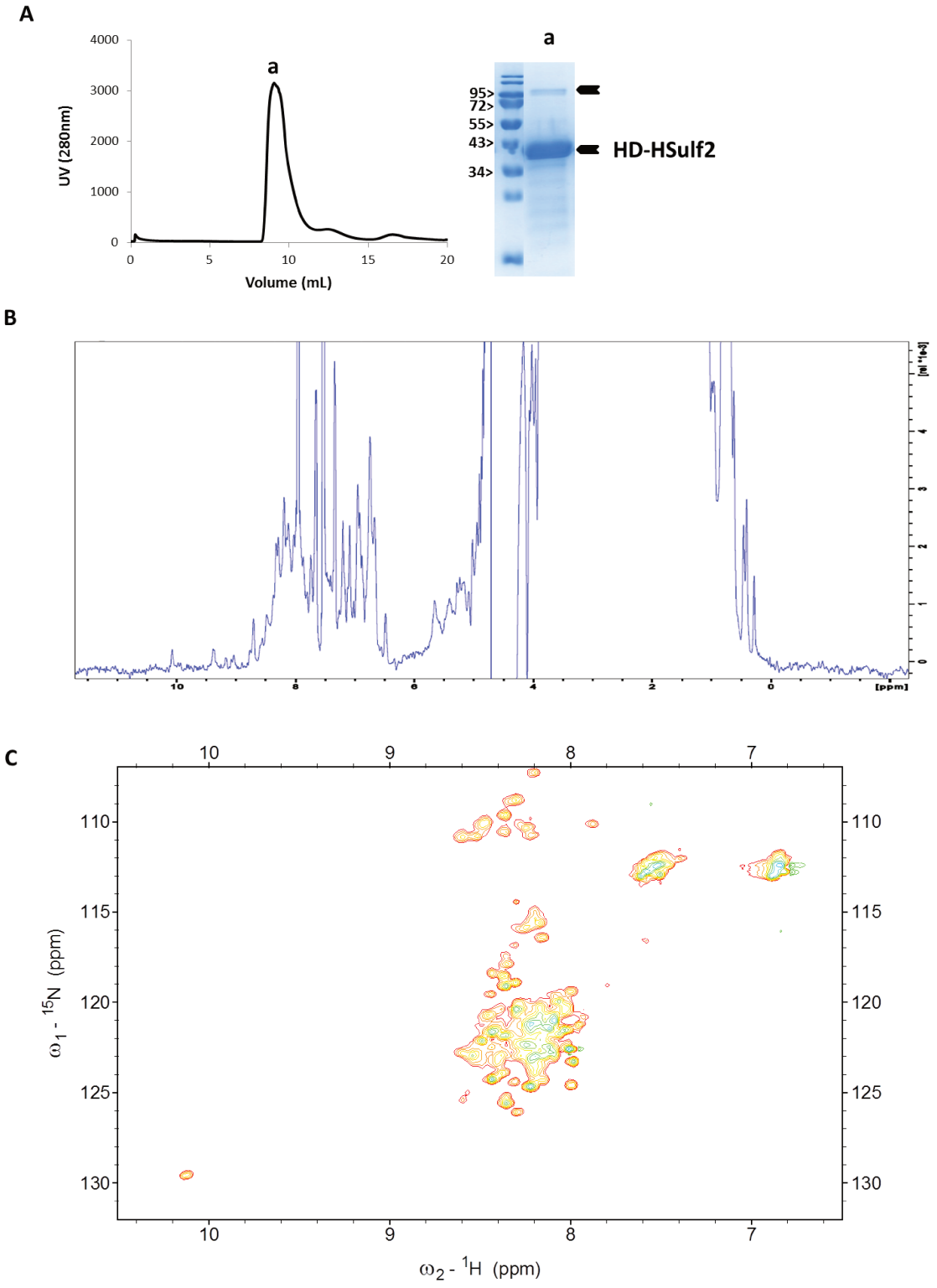


Figure 44: Expression and structural characterization of HD-HSulf2. Size-exclusion separation profile of HD-HSulf-2. Fraction corresponding to the peak a was analyzed by PAGE and Coomassie blue staining (A). NMR 1D profile of purified HD-HSulf2 (B). NMR 2D spectrum of ^{15}N labeled HD-HSulf2 (C).

2D NMR analysis of HD at 120 µM showed as well no resonances between 9 and 10 ppm and confirmed therefore that the HD is at least very dynamic (Figure 44). Very interestingly,

we observed several new resonances outside the 8 – 8.5 ppm range, with an homogeneous intensity and form, when the HD was not well conserved at 4°C and was degraded into a 25 kDa major form (HD-deg) (Figure 45). This suggests the presence of a structured region in this last form. Degradation mostly occurred in the C-terminal part of HD, as the HD-deg band could be visualized by western blot, using an anti-His antibody (data not shown - of note, His tag is at the N-terminus of the protein, when expressed with the pET28a vector).

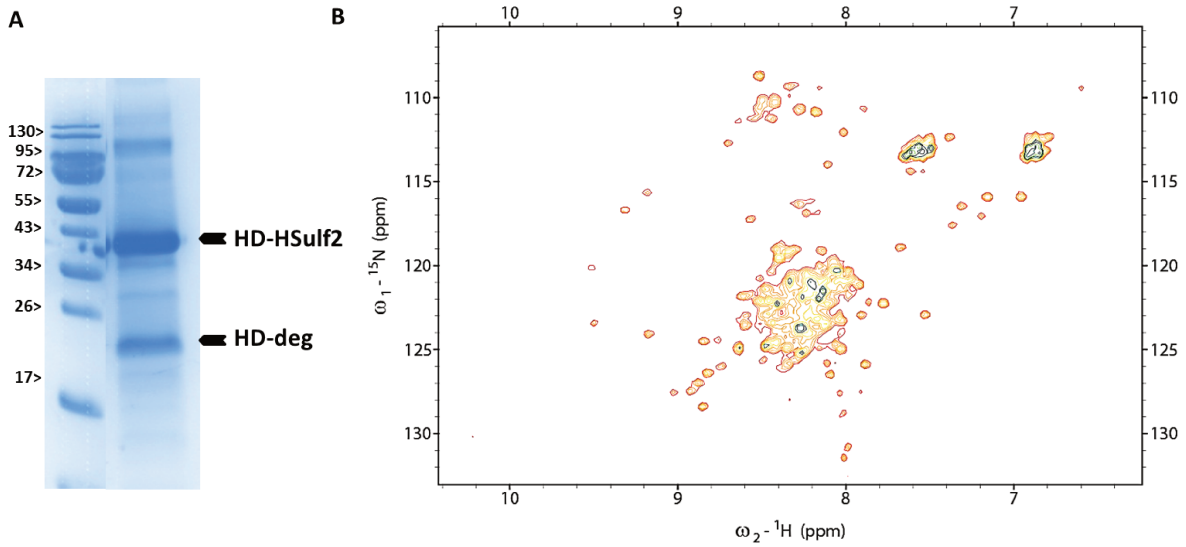


Figure 45: Structural characterization of HD-HSulf2 containing HD-deg form. SDS electrophoresis gel of HD-HSulf2 showing a degraded form of HD-HSulf2 (HD-deg) (A). NMR 2D spectrum of ^{15}N labeled HD-HSulf2, containing the degraded form (B).

We sought of adding the substrate of HD (HP) in the aim of stabilizing the unstructured regions of the HD. For this, we added on the HD a [IdoA(2S), GlcNS(6S)]$_2$ tetrasaccharide (dp4, see page 178) in a 1:1 ratio. The degree of folding of the protein did not change, but we saw some chemical shifts, pointing out interactions of some residues (Figure 46). We decided then to use NMR in order to identify HS binding sites within the HD.

First, we worked on the optimization of the NMR spectrum resolution. As slightly acidic buffers are known to improve spectrum resolution, we tested several buffers to select one yielding better-resolved peaks, without compromising protein solubility. Best results were obtained with 50 mM MES, 100 mM NaCl, pH 6, which significantly improved peak resolution and did not caused any significant protein precipitation (Figure 47).

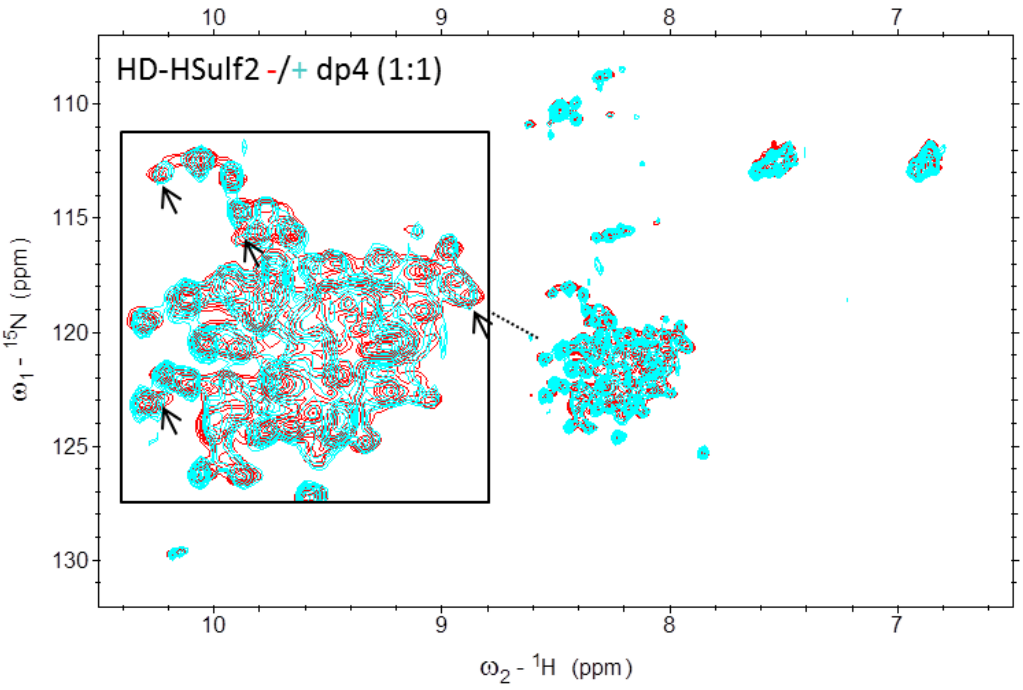


Figure 46: Interaction of HD-HSulf2 with a HP tetrasaccharide. NMR 2D spectrum of ^{15}N labeled HD-HSulf2 in the presence (blue) or not (red) of a [IdoA(2S), GlcNS(6S)]$_2$ tetrasaccharide (dp4), using a 1:1 ratio.

A significant issue to address in this part of the project was that the large size of the HD domain (37 kDa) could hinder the acquisition of spectrum with a high enough resolution to allow peak attribution. To address this and based on the fact that the HD domain is very dynamic, we speculated that HS binding sites within this domain could be studied independently. To determine the HS binding sites, we sought of applying the beads approach on the HD (Vivès et al., 2004). As a reminder, briefly, this consists on creating a covalent binding between HD and HP coated beads (see CMLS article, page 90). The unbound parts of HD are removed with proteolysis, while the bound motifs are identified by N-terminal sequencing of the beads. Results showed the presence of three basic clusters in the HD that may interact with HS: R_{518}RKKLFKK$_{525}$, R_{649}GHLKKKR$_{656}$ and K_{702}RKKKLRK$_{709}$ (Figure 48). Based on these putative HP binding sites, we designed 4 small ~ 12 kDa overlapping constructs (called HD-A, HD-B, HD-C and HD-D) spanning the entire HD sequence, in which two constructs comprised these basic clusters (B and D), but not the two others (A and C) (Figure 48).

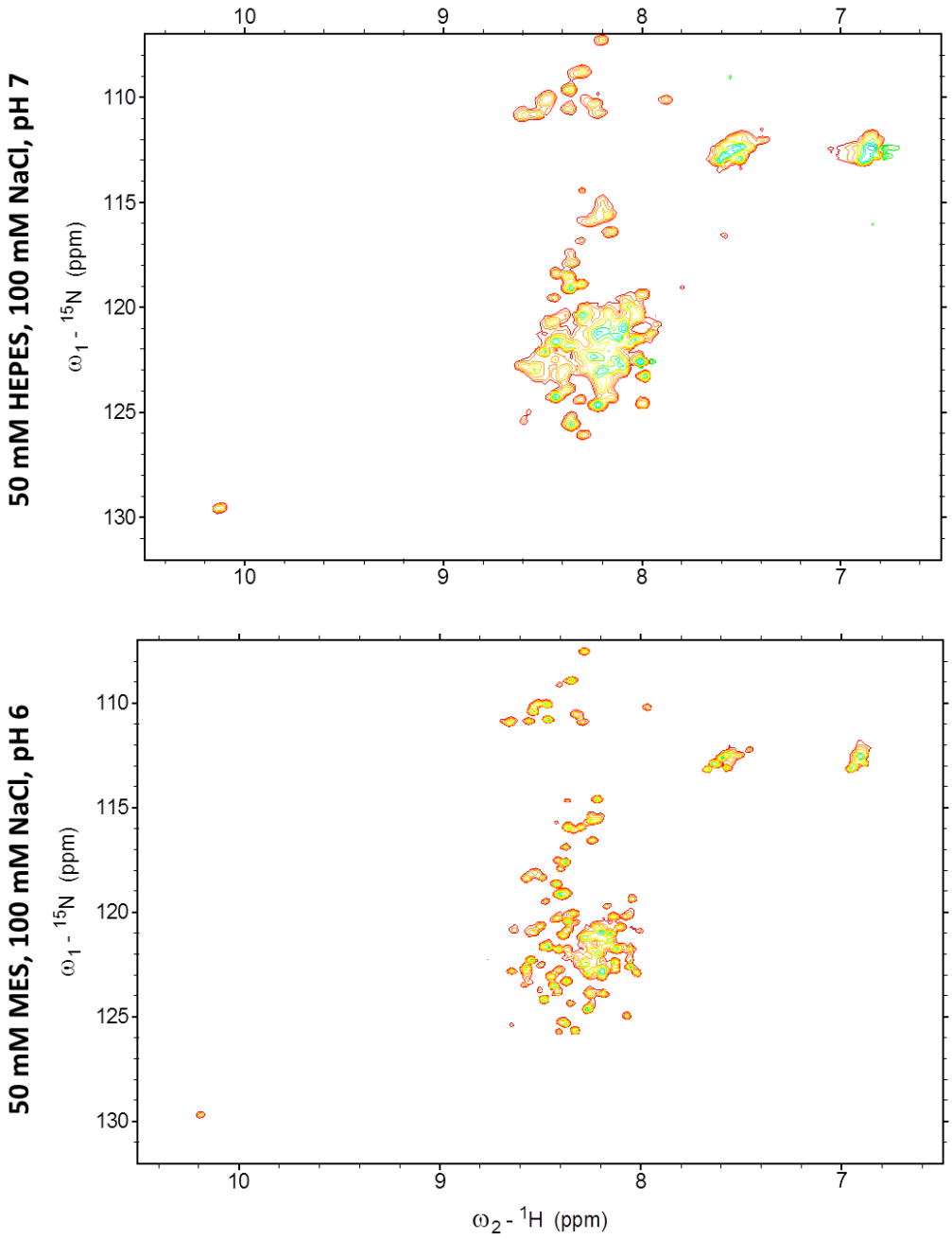


Figure 47: Comparison of the resolution of NMR 2D spectrum of HD-HSulf2 in two different buffers.

We produced and purified the HD A-D proteins, as the wild type. All the constructs showed similar size exclusion chromatography profiles (Superdex75 10/300 GL). They were eluted at 13 mL and the peak was analyzed on Coomassie blue stained PAGE, showing a pure major band at 12 kDa, which corresponds to the expected MW (Figure 48). A dimeric form may also exist, as the HD WT, given that the PAGE showed minor band at 24 kDa. Proteins were then isotopically labeled, using conditions developed for full length HD, and 2D NMR spectra were acquired (with the help of Pierre Gans, NMR IBS). All the constructs

yielded typical unstructured protein spectra, except for the HD-A, where some structured features could be detected, as already observed for the degraded HD form (Figure 48, Figure 49)

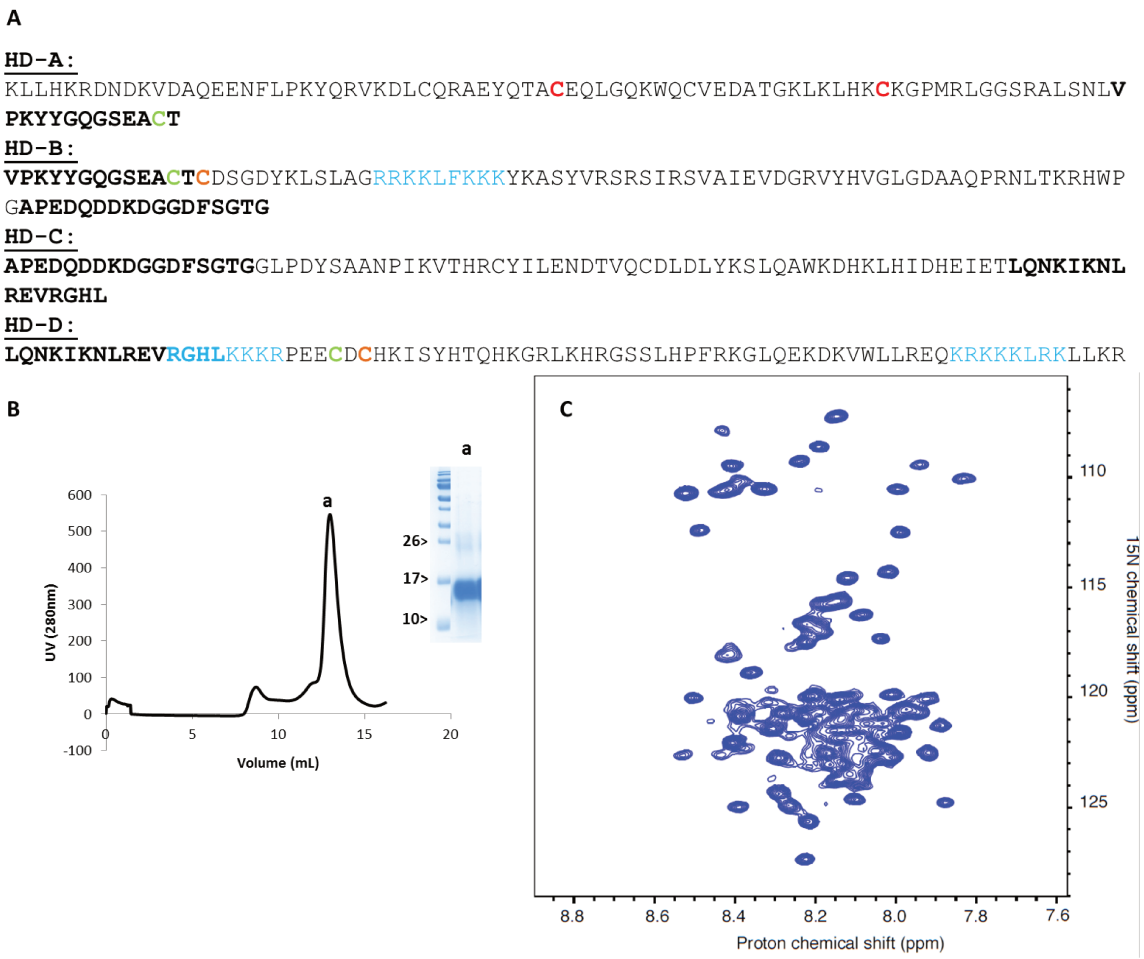


Figure 48: Expression and characterization of HD-HSulf2 constructs. Amino-acid sequence of HD-HSulf2 showing the four constructs HD-A to D. Overlapping sequences are shown in bold. The basic cluster identified using the heparin bead approach are indicated in blue. Each cysteine pairs forming a disulfide bond identified in HSulf-2 WT are highlighted with the same color (red, green and orange- see discussion, page 155) (A). Size-exclusion separation profile of the HD-B. Fraction corresponding to the peak a is analyzed by PAGE and Coomassie blue staining (B). NMR 2D spectrum of ^{15}N labeled HD-B construct (C).

We then added a dp4 [IdoA(2S), GlcNS(6S)]$_2$ (see page 178) in a 1:1 ratio prior to NMR acquisition, and we identified the constructs, for which chemical shifts could be observed, as a result of an interaction with the tetrasaccharide. For HD-C, no change was observed upon dp4 addition, strongly suggesting that HD-C and dp4 do not interact. For HD-A and HD-D, we observed shifts of some resonances, indicating interaction with dp4 of both species (Figure 49). For HD-B, the addition of dp4 resulted in an immediate precipitation of the formed complex, as observed by the complete loss of the protein signal (data not shown).

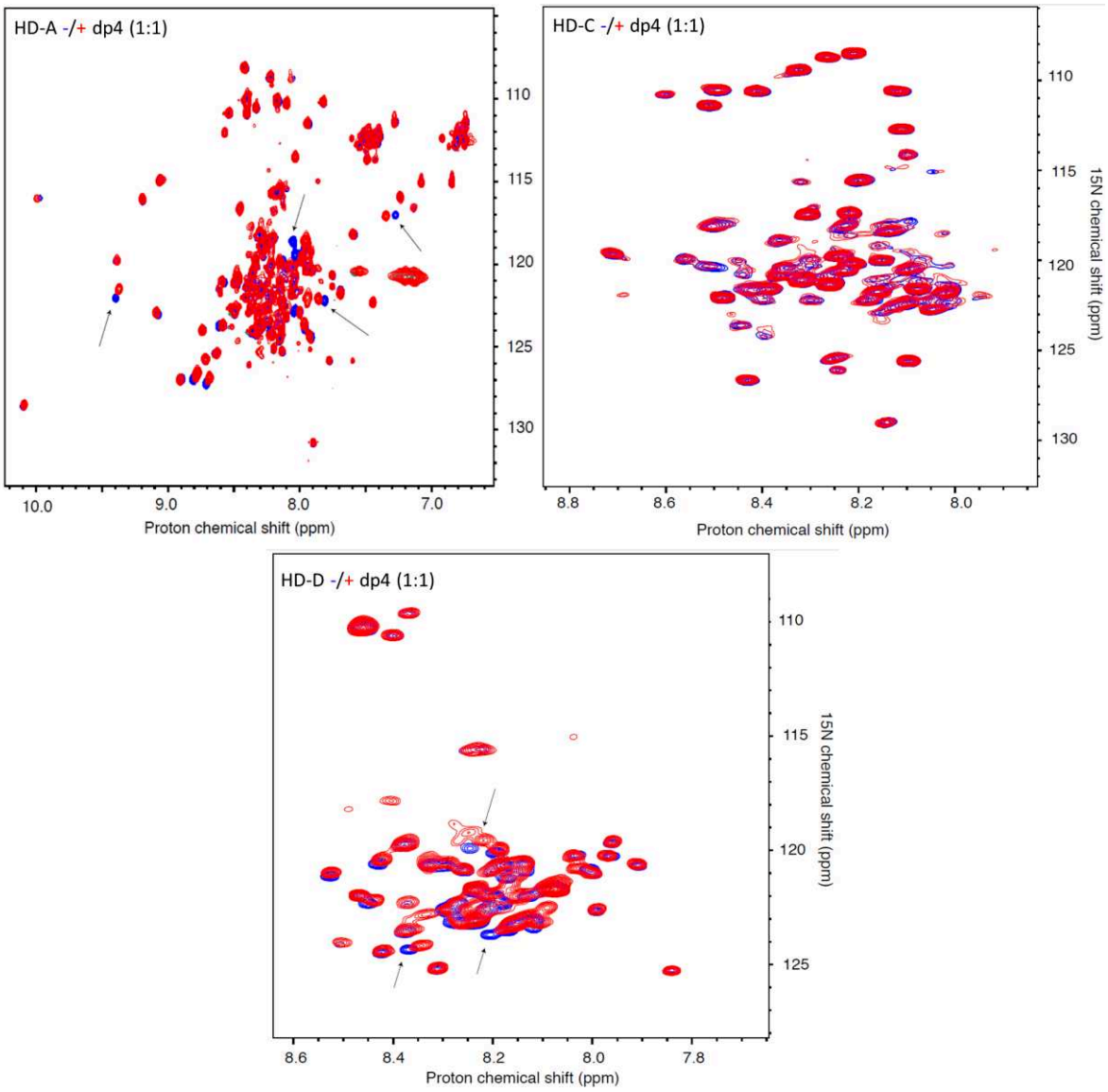


Figure 49: Interaction of HD-A, HD-C and HD-D with a HP tetrasaccharide. NMR 2D spectrum of ^{15}N labeled HD-A, HD-C and HD-D in the presence (red) or not (blue) of the [IdoA(2S), GlcNS(6S)]$_2$ tetrasaccharide (dp4), using a 1:1 ratio.

We next undertook 3D NMR study to identify the dp4 binding site in HD-A and HD-D (with the help of Pierre Gans and Alicia Vallet, NMR, IBS). The first step was to assign the resonances of both constructs. This was made using ^{13}C ^{15}N labeled samples and 3D NMR experiments (HNCA, HNCACO, HNCACB, HNCO, HNCOCACB). HD-D was fully assigned (by Yoan Monneau). Results indicate that at 10°C, HD-D is essentially unfolded, except in the C-terminal part (E_{700}-R_{713}), which shows a strong helical propensity. Interestingly, this region contains the putative HP binding motif. This work is still in progress.

III. Functional analysis of HD

The HD is produced in a prokaryotic system, lacking all the PTMs (furin cleavage, *O*-glycosylation and maybe *N*-glycosylations) and is solubilized by flash refolding. Usually, protein refolding efficiency can be estimated by its biological activity, such as enzymatic activity. However, the HD lacks any active site. It is thus important to validate the structural and functional relevance of this isolated recombinant domain.

Functionally, we sought of testing the ability of the HD to bind heparin using a set of different techniques, including ELISA and FACS. ELISA showed that HD binds to immobilized HP, and the binding increases with the concentration of HD (Figure 50). In line with this, FACS experiment showed that, as HSulf2ΔSG, HD can bind to WISH cell surface HS, although less efficiently (Figure 50). The differences may be due to the HSulf2 CAT domain, as we already showed that it is implicated in the interaction with HP (see CMLS article, page 90).

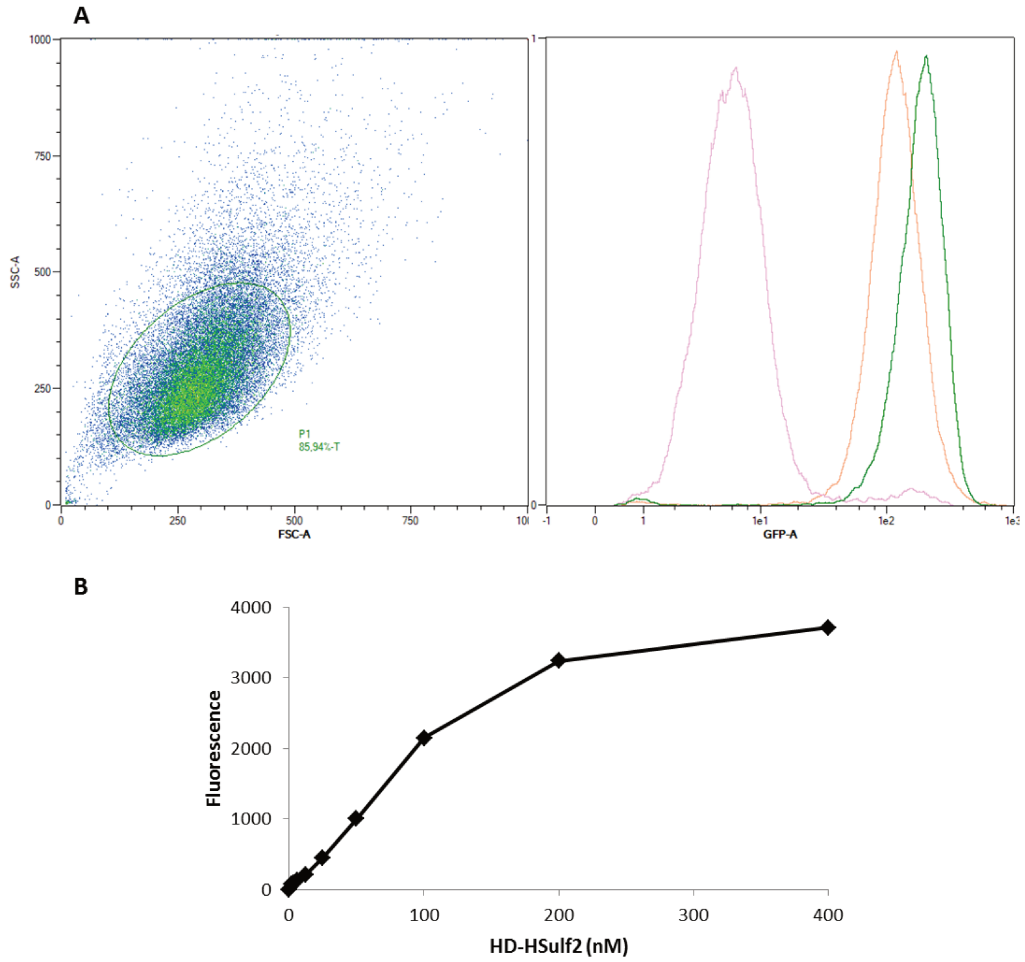


Figure 50: Studying the biological activity of HD-HSulf2. FACS analysis of HD-HSulf2 (orange) or HSulf2ΔSG (green) binding to the surface of Wish cells. The background fluorescence level of the antibody is shown in pink (A). Immunoassay of HD-HSulf2 binding to immobilized HP (B).

Structurally, we performed Circular Dichroism (CD) analysis of the HD, HSulf2ΔHD and HSulf2ΔSG and Small-angle X-ray scattering analysis (SAXS, Kratky plot) to the two latter enzymes (with the help of Julien Perard). These two techniques consist on evaluating the secondary structure and the folding properties of the protein. Our goal was to validate the presence of a destuctured region within the HSulf-2 enzyme and to confirm that these could be due to the HD domain.

CD analysis consists on measuring differences in the absorption of left handed polarized light versus right handed polarized light, which arise from structural asymmetry. Each light has its own electric field (E). The plane of the light wave is rotated and the addition of both E vectors results in one that traces out an ellipse. CD is reported thus in degree of ellipticity, which is defined as the angle whose tangent is the ratio of the minor to the major axis of the ellipse. For proteins, far UV (180-260 nm) and near UV (250-330 nm) circular dichroism measurements give insights respectively into their secondary structure content and their tertiary organization, where structural features have characteristic CD spectra profiles. CD requires to exchange the buffer to an optimal one for CD (5 mM Trizma, 100 mM NaF, pH 7.5) using several steps of dilution washes and centricon concentration. Proteins were analyzed at a concentration of 10 µM. Results showed that HSulf2ΔSG and HSulf2ΔHD have negative bands at 222 nm and 208 nm and a positive one at 193 nm (Figure 51). These spectra profiles correspond to a structured protein containing α-helices. However, HD has very low ellipticity above 210 nm and negative bands near 195 nm, indicating a disordered protein (random coils structure).

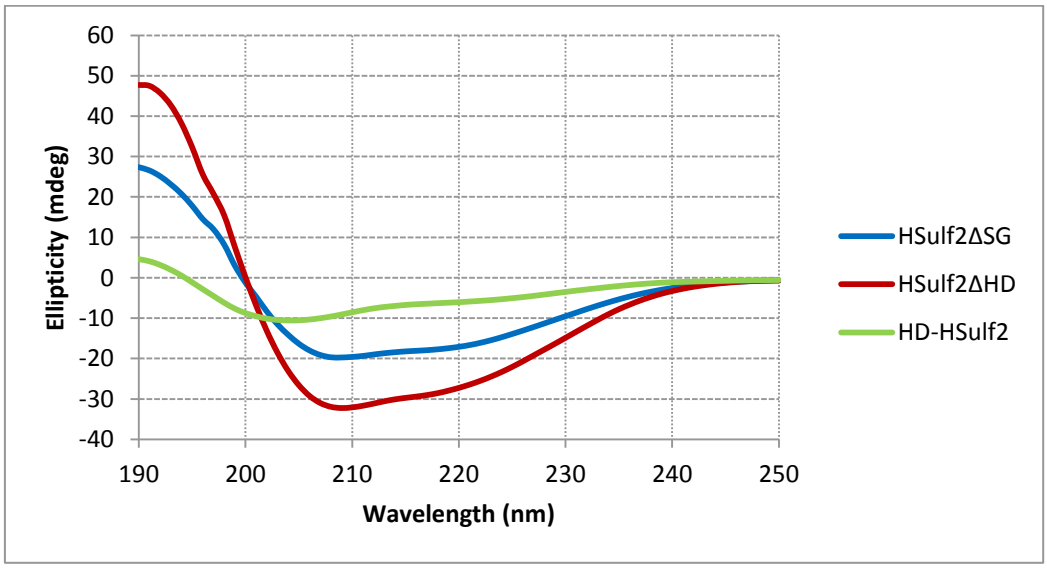


Figure 51: CD UV spectra for HSulf2ΔSG, HSulf2ΔHD and HD-HSulf2.

SAXS is a biophysical method to study the overall shape at low resolution and structural transitions of biological macromolecules in solution. The Kratky representation helps visualizing particular features of the scattering profiles for an easier identification of the folding state and flexibility. As Figure 52 shows, HSulf2ΔHD represents a bell-shape distinct peak, indicating a globular protein. Regarding HSulf2ΔSG, it represents a peak accompanied with a rise to a plateau, suggesting a partially folded protein with an elongated portion.

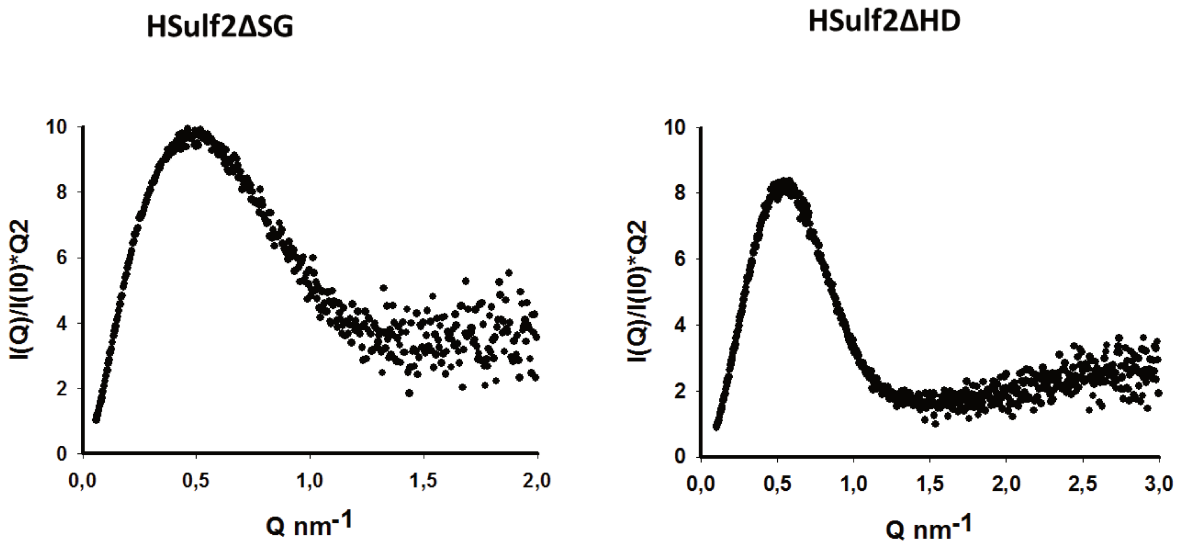


Figure 52: SAXS Kratky plots for HSulf2ΔSG and HSulf2ΔHD.

These results suggest that the unfolded regions found in HSulf-2 could be mostly attributed to the HD domain. In line with this, HD is very sensitive to limited trypsin digestion performed on HSulf2ΔSG.

CHAPTER

VI

Discussion and perspectives

GAGs, including HS, are complex sulfated polysaccharides found abundantly at the cell surface and in the ECM. They can bind to most signaling proteins and control their activity in order to trigger some specific cell functions. HS/ligand binding depends on highly sulfated saccharides regions called the NS domains. To allow rapid adaptation of the cell to changes in its environment, HS chains are rapidly recycled (half-life of ~ 3-4 hours) and renewed at the cell surface with different chain length and variable sulfation. The structural organization of HS is controlled by a complex machinery of biosynthesis enzymes, and also by post synthetic enzymes including heparanases, sheddases and sulfatases of the Sulfs family.

Sulfs are extracellular endosulfatases that remove specifically the 6-*O*-S groups found within HS NS domains. By targeting these groups, Sulfs significantly affect the polysaccharide biological activities, and have thus been implicated in various physiopathological processes, notably cancer. The aim of the project was to characterize the functional and structural properties of HSulf-2 isoform.

1. Production of recombinant Sulfs

Studying Sulfs was hampered by the limited access to recombinant enzyme. Recently, we have unlocked this critical bottleneck by setting up a new protocol for the expression and purification of recombinant Sulfs. This consisted in the expression of these enzymes in mammalian cells and not in bacteria, since their activity depends upon important PTMs, including the cysteine converted formylglycine residue, the furin cleavage and *N*-glycosylations. We chose to produce Sulfs in suspension-adapted cells (HEK293F) using serum free medium, with SNAP and His tags at their N- and C-terminus respectively, to facilitate detection and purification. This system was validated by confirming enzyme activities of produced recombinant HSulf-2, e.g. their ability to desulfate aryl compound 4MUS, and to bind and to digest their natural substrate HP and HS *in vitro*. We also performed *in cellula* characterization assays on HSulf-2. We confirmed the ability of recombinant HSulf-2 to bind to cell surface HS using FACS assay, and initiated a study to investigate its role in modulating SDFα chemokine activity. Preliminary results (not shown in this manuscript) suggested that recombinant HSulf-2 reduced the capacity of SDFα to trigger cell signaling and cell migration. For recall, previous studies from the lab had also demonstrated the ability of recombinant HSulf-2 (as concentrated conditioned medium, not as a purified protein) to tune HS-mediated FGF cell proliferation (Seffouh et al., 2013). Finally, as part of collaboration with Eric Schmidt team (university of Colorado), we showed that purified recombinant HSulf-1 could also be used *in vivo* (Oshima et al., in revision). In this study, we tested the effect of the recombinant enzyme on post-septic lung mice. Sepsis is an hyper inflammatory process, followed by a delayed period of immunosuppression called compensatory anti-inflammatory response syndrome (CARS), which can lead to secondary inflammation. It has been shown by Eric Schmidt team that HS of lung endothelial glycocalyx displayed higher 6-*O*-sulfation content after septic injury, which was due to downregulation of Sulf-1. Interestingly, they showed that post-septic loss of Sulf-1 was necessary for CARS to occur, and that administration of exogenous recombinant Sulf-1 intravenously reversed the immunosuppression phenotype. These data underline the efficiency of our expression system and open new perspectives for using recombinant Sulfs in elaborate biological models to study their roles in various physiopathological processes.

Post-translational modifications of HSulf-2

Access to purified enzyme allowed us to investigate the PTMs of HSulf-2. For this, we collaborated with Regis Daniel team (University of Evry) to perform a mass spectrometry (MS) study of HSulf-2 (Seffouh et al., 2019). Results showed a MW of 133.15 kDa for the enzyme, highlighting the presence of 35 kDa of PTMs (the theoretical MW of HSulf-2 is 98.15 kDa). The catalytic formylglycine was identified at the cysteine 88 position. The two catalytic signatures, identified by MS, which are well conserved amongst sulfatases, were located from C_{88} to G_{98} and from G_{135} to K_{143} for HSulf-2. MS analysis estimated that 10.23 kDa of PMTs corresponded to *N*-glycosylations and showed the presence of *N*-glycans on four aspargines (Asn, N) out of the seven potential sites in the CAT domain: N_{132}, N_{171}, N_{168} and N_{241} (see page 174). However, it is important to note that glycosylations may vary among cell lines. The remaining 24.72 kDa may thus correspond to the GAG chain, which is composed of alternating sub-domains of CS (CS-A and/or CS-C) and DS (CS-B).

Regarding furin processing, 7 potential cleavage sites have been proposed, based on amino-acid sequence prediction (Morimoto-Tomita et al., 2002), and 2 were experimentally validated by N-terminal sequencing (Tang and Rosen, 2009). Edman sequencing of our recombinant purified protein showed only cleavage after the R_{533}–R_{538} site. However, western blot analysis of the C-terminal part of HSulf2ΔSG showed two bands at close MW, which suggests the presence of two furin cleavages in the HD domain, presumably at the sites determined by Tang and colleagues: the R_{533}–R_{538} primary site (always detected) and a R_{560}–R_{565} secondary site, on which occasional cleavage may occur. In this context, it would be interesting to investigate whether this secondary cleavage that we only observed on the HSulf2ΔSG could be impaired by the presence of the GAG chain, which is anchored at the vicinity of the secondary site. Heparin binding to proteins like HIV-1 gp160 (Pasquato et al., 2007) or semaphorin 3A (Djerbal et al., in preparation) has been shown to enhance furin cleavage, which may contradict this hypothesis. In addition, it is not clear when the furin cleavage occurs during the synthesis and/or post-translational processing of the enzyme. Furin cleavage has been reported to take place in the Golgi and in line with this, we could detect the cleaved form in the cell lysate (of note, this includes both Golgi and cell surface bound proteins). However, our data also suggest that not all the secreted enzymes are cleaved by furin, given the presence of the unmaturated form in the extracellular medium.

Regarding the disulfide bonds, it is very important for us to identify their presence, whether they are intra or inter domain (between CAT and HD domains), especially for our

structural study. Preliminary results provided by our collaborators, showed the presence of three intra HD domain disulfide bonds, two of them bridging the N-terminal and C-terminal subunits of the enzyme (of note, the HD domain spans over both sub-units). This is an encouraging structural observation, regarding our HSulf2ΔHD construct, and more importantly regarding the HD domain. The disulfide bonds involve C_{455} and C_{477}, C_{504} and C_{660}, C_{506} and C_{662}. The first one is located before the furin cleavage site, whereas the others link the two parts of HD after the furin cleavage (see page 174).

MS analyses are still in progress to validate all the *N*-glycosylation sites, to confirm the disulfide bonds and to identify the furin cleavage sites.

2. 6-*O*-desulfation mechanism of HSulf-2

Studies in the lab have previously shown that HSulfs catalyzed the 6-*O*-desulfation of HS through an original process and oriented mechanism (Seffouh et al., 2013). Access to purified Sulfs then allowed us to further analyze the HS/Sulfs recognition processes. It has been well established that HSulf-2 CAT domain comprises the enzyme active site, that the HD is responsible for the high affinity binding to HS, and that the digestion of HS requires both domains. Here, we showed that HSulf2ΔHD could also interact with HS, thanks to two newly identified binding sites: V_{179}KEK and L_{401}KKK. These two epitopes align with the active site located within a groove in between, and may participate to the activity by guiding and enabling proper presentation of the substrate to the catalytic FGly residue (see CMLS article, page 90). We also showed that the HSulf2ΔHD exhibit an exosulfatase-like activity on heparin oligosaccharides, suggesting that the HD domain is not strictly required for HS recognition, but is essential for efficient and processive desulfation of the polysaccharide (Figure 31).

In parallel, we investigated further HSulf-2 substrate specificity. First, the minimal oligosaccharide size required for the enzyme function was determined to be a tetrasaccharide (Figure 31). Second, we showed that presence of the terminal unsaturated uronic acid was not essential for the enzymatic activity (Figure 32). This last result suggests that a trisaccharide devoid of non-reducing terminal uronic acid could be the actual minimal substrate usable by the enzyme. To test this, we are planning to generate such trisaccharide, by treating our tetrasaccharide with mercuric acetate (Vivès et al., 1999). Noteworthy, the distance that spans the two VKEK and LKKK sites approximately corresponds to a dp8 length, thus suggesting that these sites may be required for processing of long HS chains but not for small oligosaccharides. In line with this, substitution of both VKEK and LKKK sites did not affect the capacity of HSulf-2 to digest octasaccharide or shorter oligosaccharides (results not shown).

Moreover, we have had access to a panel of synthetic octasaccharides differing by the position of one missing 6-*O*-S. We showed that HSulf-2 could digest all these octasaccharides. However, we could not test the octasaccharide missing a 6-*O*-S at position 2 (Figure 33). It would be important to see whether HSulf-2 would be able to reach the third 6-*O*-S despite the absence of the second one. This could help us to better understand the catalytic mechanism of HSulf-2 *in vivo*. Indeed, HP octasaccharides used in this study are

homogenous and fully sulfated oligosaccharides, but such structures are rarely found in physiological HS. Synthetic octasaccharides could thus represent more relevant mimics of HS NS domain structures. In that line, it would be interesting to have access to synthetic oligosaccharides missing internal *N*- or 2-*O*-sulfates to test the ability of HSulf-2 to use such incompletely sulfated structures as substrates.

3. HSulf-2 as a CSPG and its implication in cancer

Access to purified enzyme also allowed us to show that HSulf-2 is a unique CSPG, bearing one polysaccharide chain anchored to its HD domain. Evidence of this was first seen during the size exclusion chromatography purification step of HSulf-2, which showed the presence of two peaks, one with an early elution time (peak 1), corresponding to a very high apparent MW (~1000kDa), and one at the expected elution time (peak 2), with a calculated apparent MW of ~150 kDa (see article page 112). Negative staining EM analysis of the first peak fractions demonstrated that it did not correspond to protein aggregates. This protein preparation was on the contrary homogenous and showed small ring shaped small particles of an estimated 150-300 kDa MW, which was in clear disagreement with size exclusion chromatography elution times. We therefore hypothesized that the enzyme could have been purified in complex with HS chains. However, treatment of HSulf-2 with heparinases did not affect size exclusion chromatography elution time. In contrast, treatment with chondroitinase ABC yielded an additional third peak (peak*) with a significantly delayed elution time compared to peak 1, but a still earlier elution time compared to peak 2 (see article page 112). The presence of this covalently linked CS chains was eventually confirmed using a number of biochemical assays (dissociation of potentially non-covalent complexes using NaCl, or urea treatment, inhibition of GAG synthesis with xylosides..).

Regarding proteins from the peak 2 size exclusion chromatography profile, we demonstrated that it corresponded to a degraded form of HSulf-2 that had lost the part of the HD domain bearing the CS chain. We thus speculated that the GAG chain is anchored to the HD domain at the level of an unstructured, highly sensitive to proteolysis loop. In line with this, addition of anti-proteases throughout the entire purification procedure performed at 4°C resulted in the disappearance of peak 2.

To understand the biological role of this CS chain, we generated a mutant of HSulf-2 that lacks the CS chain (HSulf2ΔSG). As expected, HSulf2ΔSG is eluted at peak* on size exclusion chromatography. We next performed *in vitro* and *in vivo* studies order to compare HSulf-2 WT and HSulf2ΔSG activities. We first showed *in vitro* that the mutant form was more active in our endosulfatase assay (ability to 6-*O*-desulfate heparin) and it bound to HS with a higher affinity (ELISA). We also showed increased binding to the cell surface using FACS analysis (see article page 112).

To study the activity *in vivo*, we chose to focus on the extensively reported role of HSulf-2 during tumor progression. In collaboration with Odile Filhol-Cochet (BIG-CEA, Grenoble), we setup an *in vivo* tumor xenograft model, based on the injection of HSulf-2 WT or HSulf2ΔSG stably transfected cancer cells in mice. To do this, we selected MDA-MB-231 cells for many reasons. These are epithelial breast cancer cells in an aggressive stage and their manipulation (injection in the mammary gland + measurements of tumors size) is easy. In addition, this model does not require exogenously added estrogens for xenograft growth. Finally and most importantly, MDA-MB-231 cells do not express HSulfs, which thus excludes unwanted effects due to endogenous forms of these enzymes.

Noteworthy, this *in vivo* model has been previously used in two precedent studies on the role of HSulf-2 during tumor progression, which reported contradictory results: pro-oncogenic (Zhu et al., 2016) and anti-oncogenic activities (Peterson et al., 2010). Our results did not show any significant pro-oncogenic activity of HSulf-2. This clear discrepancy could be explained by differences in the experimental conditions between the three studies. We analyzed the effect of HSulf-2 until the tumors reached 1 cm^3, which is the limit defined by European ethical rules. In the Zhu *et al.* study, the tumors were allowed to reach 3-4 cm^3 before sacrifice of the animals. We thus hypothesize that the absence of significant difference between control and HSulf-2 transfected conditions in our *in vivo* experiment could be due to an interruption of the experiment at a too early stage. Regarding the Peterson *et al*. study, the effect of HSulf-2 expression was analyzed on 100x smaller tumors (0.04cm3). Furthermore, HSulf-2 overexpression in MDA-MB-231 cells was analyzed at the mRNA but not the protein level. This may be not reliable enough, in light of our first transfection trials, during which expression of Sulf was lost after several passages. However, we cannot exclude that HSulf-2 could have different effects depending on the cancer development stage. Furthermore, in the same study, the injection of recombinant HSulf-2 in tumors did not affect their size. Noteworthy, when we strived to establish stably transfected cells, we tried to inject our recombinant HSulf-2 in tumors. Our preliminary results suggested a decrease of the tumor size in the presence of HSulf-2. These data underlined the importance of the microenvironment on HSulf-2 effect. Indeed, the HSulf-2 stably producing cells system mimics more the physiological context than the injection of recombinant protein.

Second, our data showed that the mutant HSulf2ΔSG significantly increased tumor growth, tumor vascularization, and invasion of other tissues (see article page 112). This is in line with our *in vitro* results, which showed increased activity of the mutant form. Interestingly, the

stability of Sulf expression was investigated in tumors after dissection. Results showed that HSulfs (WT or mutant) could be detected in all tumors (based on the whole length enzyme). Western blot analysis using an anti- C-terminal HSulf-2 antibody revealed that HSulf-2 WT had lost its CS chain. We hypothesized that this could be due to proteolysis of the HD domain by extracellular proteases, as was observed for the peak 2 fraction during the purification of the recombinant enzyme. Natural processing of the HD *in vivo* could thus lead to a heterogeneous population of enzyme with or without GAG chain over time, which may explain the wide distribution of the data obtained in our *in vivo* assay for tumors of the HSulf-2 WT conditions, as well as the clear pro-oncogenic activity of HSulf-2 observed by Zhu and colleagues in large, late-stage tumors.

Indeed, in pathological conditions like cancer, the expression of matrix metalloproteinases (MMP) is induced by cytokines including as TNFα or IL1β, in order to degrade extracellular matrix proteins and trigger cell functions such as proliferation, angiogenesis, and differentiation (Gialeli et al., 2011; Itoh and Nagase, 2002). Proteoglycans are amongst these targeted extracellular proteins. For example, the extracellular domain of syndecans bearing the GAG chains, is cleaved from the rest of the protein by a process called shedding.

In line with our data, we thus propose that HSulf-2 GAG chain is a major regulatory component of the enzyme activity, which limit access to HS substrate. This prevents the anchoring of HSulf-2 on cell surface, resulting in the diffusion of the enzyme.

However, in specific conditions such as within the tumoral microenvironment, induced extracellular proteases may catalyze the cleavage of the HD domain region bearing GAG chain, resulting in an “activation” of the enzyme and an increase of its pro-oncogenic functions. HSulf-2 could then bind HS with higher affinity in the ECM, and 6-*O*-desulfate the polysaccharide with higher efficiency, leading to the release of sequestered signaling proteins, such as pro-oncogenic growth factors and cytokines, which could in turn induce proliferation, migration and angiogenesis (Figure 53).

By analogy with syndecan shedding process, we also hypothesize that HSulf-2 CS chain could play other important roles, such as the recruitment of protein partners to trigger specific functions, and that proteolysis released fragments containing the CS chain could be involved in other biological functions.

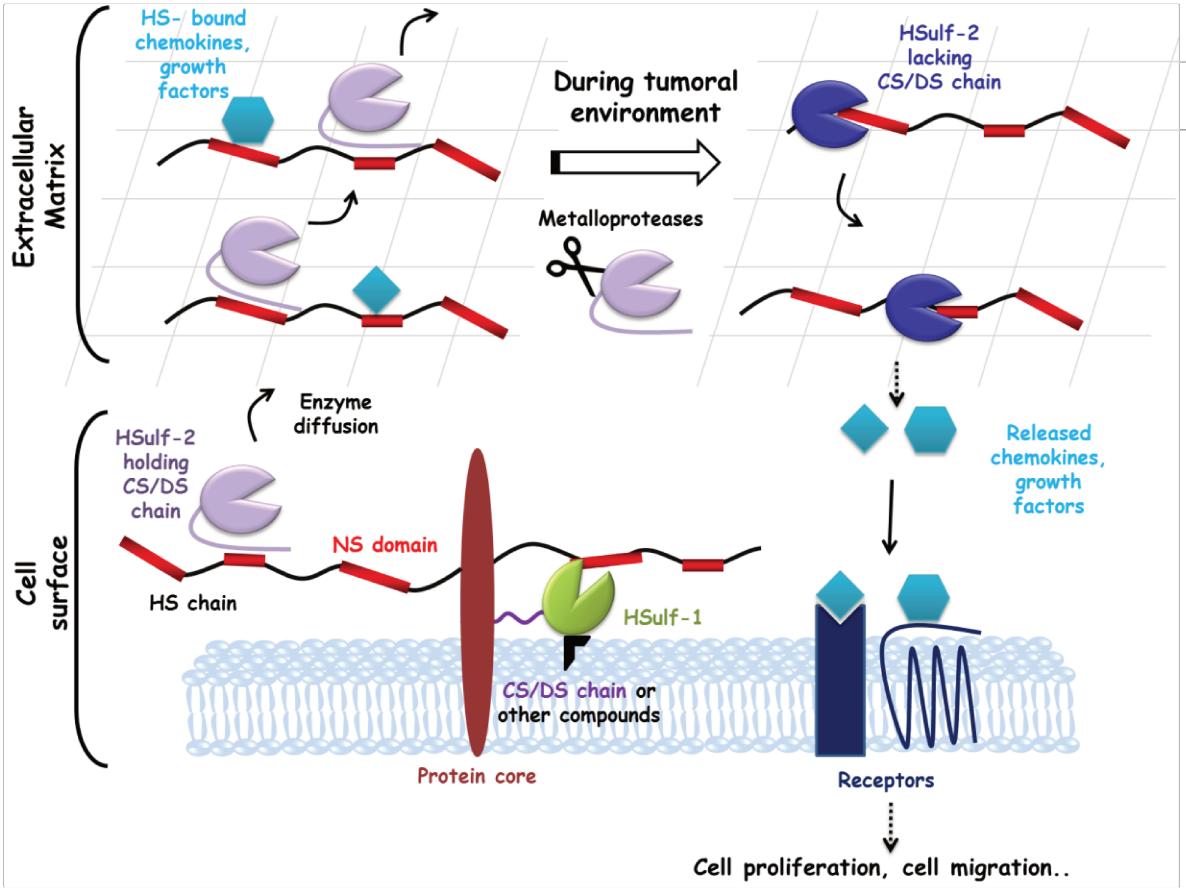


Figure 53: Model depicting the mode of action of HSulf-1 and HSulf-2 in a tumoral microenvironment.

To our knowledge, HSulf-2 is the first enzyme exhibiting a GAG chain. This novel discovery opens completely new perspectives in the study of the Sulfs. In the future, it will be critical to determine if the GAG chain is present in all the physiological conditions and in all Sulf-2 expressing cells. For example, PGs like serglycin exhibit CS or HP chains depending on the cell type. In addition, some PGs do not always harbor their GAG chain. Our results showed that the CS chain is present on the enzyme in two different overexpression systems (HEK293F and MDA-MB-231 cells). In addition, preliminary results showed that MCF7 cells, another invasive breast cancer cell line, express HSulf-2 endogenously as a CSPG (data not shown). We may also wonder what is the role of the CS chain, and of the putative HD proteolysis mechanism releasing this CS chains, in other disease or physiological processes, in which HSulf-2 is involved. In fact, Sulfs are known to play important roles during development, by regulating the activity of many morphogens activities. We suggest that the

distribution of Sulfs among tissues may thus be dependent on the presence or not of the CS chain, which may be implicated in the development regulatory processes.

Finally, it will be interested to investigate the presence of the CS chain in other orthologs such as *Drosophilia* or *C. elegans.* Our results, following the expression and purification of Sul1 (*C. elegans* form), suggest that Sul1 did not represent a CSPG.

Analyses are in progress in order to characterize the CS chain: the exact type, the composition and the sulfation level.

4. Structural analysis of HSulf-2

One of the main objectives of my PhD was to undergo structural studies of the Sulfs using X-Ray crystallography. Unfortunately, crystallization assays performed on HSulf-2 failed to provide protein crystals. This could be due to the *N*- and *O*-glycosylations of the enzyme and to the highly flexible HD domain. Based on our understanding of the protein structural properties (the *N*- and *O*-glycosylations of the enzyme and the highly flexible HD domain), we pursued different strategies to obtain crystals.

Given that the CAT is common to all other sulfatases, including previously crystallized ones, and that the HD is suggested to be flexible, we sought of producing the HSulf2ΔHD form, as performed in previous published articles for HSulf1ΔHD (Frese et al., 2009). The HD sequence was removed from HSulf-2 by site directed mutagenesis, and the C-terminal domain was directly attached to the CAT domain. Interestingly, production rates of this constructs were significantly improved compared to those of the whole length enzyme. This may be due to fact that the HD binds to cell surface through HS or other components, and thus that its deletion increases the release of the enzyme in the extracellular medium. However, NaCl treatment of the HSulf-2 stably transfected HEK293F cells led to the release of only little amounts of HSulf-2 (data not shown).

We next sought of removing the SNAP tag from HSulf2ΔHD. SNAP is a large 23kDa tag, which facilitates the visualization of the enzymes in the extracellular medium, using specific SNAP binding fluorescent substrates. However, even though the SNAP itself is structured (as seen by NMR, data not shown), it is linked to Sulf through a flexible loop and could thus hinder the formation of crystals. We then attempted crystallization trials in presence of sulfocalixarene (Figure 36), a compound used in structural studies to increase the stabilization of proteins (McGovern et al., 2012). In addition, we produced HSulf2ΔHD with a reduced level of glycosylation in order to decrease heterogeneity that could prevent the crystallization.

Finally, another strategy that we developed to remove as many as possible of unstructured regions that could prevent crystallization, was to perform trypsin controlled digestion of the protein. Results showed that trypsin mostly degraded the enzyme HD, confirming the (partially or fully) unstructured nature of this domain (Figure 39). Our preliminary experiments also indicated that the first N-terminal residues of the CAT domain were also removed, suggesting that they are not structured. Interestingly, we observed that following

trypsin digestion of HSulf-2, both CAT and C-terminal domains co-eluted in size-exclusion chromatography, indicating that these domain interact together within the whole protein.

In line with this, the CAT domain alone of HSulf-2 was also produced (lacking the C-terminal region of the protein, data not shown). However, expression rates were very poor, suggesting an important role of the C-terminal in the production/stability of the CAT domain. This again supports our above mentioned results indicating interactions between CAT and C-terminal domains. Unfortunately, all the crystallization assays we tested failed. However, we could deduce important structural information that open the way to new perspectives.

First, although HSulf2ΔHD construct is active on a 4MUS substrate, and showed a well folded protein in CD and SAXS experiment (Figure 51, Figure 52), we cannot rule out that the removal of HD may affect the native structure of the enzyme. Optimization of this construction, by adding linkers between the CAT and C-terminal domains, is in progress. Second, trypsin cleaved sequences have not been precisely determined yet, but this information could be critical for the design of new optimized constructs for further crystallization assays. Third, the enzyme and especially the HD domain, has a high plasticity and therefore co-crystallization of the enzyme with its ligand (HS or HP) could stabilize its flexible regions. The complex, however, must be stable enough to yield diffracting crystals. The interaction of HSulf-2 with HP oligosaccharides could therefore improve the quality of the crystallization. However, there is a risk that the enzyme digests the oligosaccharide during the crystallization process, leading to further heterogeneity and/or dissociation of the complex. To avoid this, we propose two different strategies. The first one would be to perform HP oligosaccharides /HSulf-2 co-crystallization assays at 4°C in order to inhibit enzyme activity. Alternatively, we could use an inactive form of HSulf-2 containing a substitution of the cysteine 88. The enzyme will thereby lack the conversion of the cysteine to FGly, which is responsible for its catalytic activity.

Finally, we can go back to our basic purification protocol to perform some optimizations. In fact, we suspect molecular heterogeneity within our Sulf preparations. Purified proteins are active, but activity levels are variable from one preparation to another. One hypothesis is that the cysteine/FGly conversion do not occurs uniformly in all secreted enzymes. To assess this, we co-expressed HSulf-2 with SUMF1 enzyme that catalyzes this conversion in eukaryotic cells. The activity of the SUMF1 co-expressed enzyme was found to be two times higher than that of HSulf-2 expressed alone. Furthermore, the 260/280 absorbance ratio is not constant in all preparation, suggesting DNA contamination for some preparations. As Sulfs are HS

binding proteins, it is most likely that even small amounts of DNA contaminants could induce protein aggregation. This is actually a major issue for the wild-type CS bearing HSulf-2, as this form elutes at high apparent MW on size exclusion chromatography, thereby preventing efficient separation from aggregates. Negative staining EM analysis of all the different chromatography fractions comprised in HSulf-2 peak 1 species indeed showed high aggregate content in the early elution times, and more homogeneous and monodisperse proteins for the late elution times (data not shown). Noteworthy, the HSulf2ΔSG form, which is most likely the protein of choice to pursue crystallization trials, elutes well after protein aggregates in size exclusion chromatography, and negative staining EM analysis confirmed the absence of aggregates in these protein preparations (data not shown). Nevertheless, optimizations of our purification protocol, including addition of a DNAse treatment step, may still be needed to achieve quality of protein preparations required for structural analysis.

Given that the HD domain is predicted to be unstructured, we decided to produce it as an isolated domain, using a prokaryotic system that would allow isotopic labeling for NMR studies. HD expressed in bacteria (Ril) was mostly found in inclusion bodies (Figure 41). This is in agreement with the literature and work from Thomas Dierks groups (University of Bielefeld, Germany), where all the experiments were performed on HD produced in fusion with a soluble protein (Frese et al., 2009; Harder et al., 2015; Milz et al., 2013; Walhorn et al., 2018). We tried this strategy by expressing HD in fusions with the MBP protein. This led to very good levels and expression as a soluble form, but once we removed the MBP (thanks to a TEV cleavage site between HD and MBP), the HD precipitated, suggesting that it was not expressed in a stable form in presence of the tag (Figure 42). To avoid this, we have developed an expression and purification protocol, based on the recovery of HD from inclusion bodies and flash refolding. NMR analysis showed that our produced HD was as expected mostly unstructured, but could include some secondary structures in its N- and the C-terminus parts (HD-A and HD-D). Surprisingly, these structural features were not seen in the whole length HD, but only in the degraded form or in smaller constructs (Figure 45, Figure 49). This may be explained by the fact that unstructured regions give rise to high intensity NMR signals compared to structured domains. Alternatively, we can also hypothesize that these structured domains may interact with other residues within the HD, in a highly dynamic way, which may hinder their detection. It could be important to assess whether these two structures extremities of the HD domain, in the presence or not of an oligosaccharide, can interact with each other to fold the whole HD domain as a large loop-

like structure. In line with this, the recently identified disulfides bonds seemed to bridge these two extremities (see page 174).

Surprisingly, HD elution time on size exclusion chromatography did not correspond to MW of the monomeric form. In addition, SDS PAGE analysis of the HD showed the presence of a minor band at ~ >= 75 kDa (Figure 44). Interestingly, SAXS and MALLS experiments performed on HSulf2ΔSG showed the presence of the protein as a dimeric form (~ 240 kDa in the absence of SNAP, data not shown). Because of the CS chain, which confers to the HSulf-2 WT a very high aMW, it is difficult to conclude about the oligomerization status of the full length form. However, the degraded form of HSulf-2 (peak 2 fraction), that lacks the CS chain, shows in MALLS the MW of a monomeric form (~ 140 kDa in the presence of SNAP). It is not clear whether the dimerization of HD is also present in physiological conditions, and what would be the structural features involved. According the literature, mammalian Sulfs contain in their HD domain, coiled-coils conserved sequence (A_{623} - E_{658}) that serve as a multimerization elements (Morimoto-Tomita et al., 2002). In addition, it has been suggested that multimerization could be due to the 50 kDa C-terminal sub-unit and is facilitated by the absence of furin cleavage (Tang and Rosen, 2009). If dimerization was proved to be physiologically relevant, it would be of interest to investigate what would be the consequences for the enzyme activity, and whether the presence of the CS chain modulates this process.

One critical concern for this part of the project was to assess whether our produced isolated HD domain was a reliable mimic of the one present within the whole protein. To test this, we compared the binding of HD and full length HSulf-2 to HP by ELISA, or to the cell surface HS of wish cells by FACS (Figure 50). HD bound to both GAGs, although with a lower affinity than HSulf-2. This could be due to the lack of furin cleavage that may contribute to high affinity binding, or to the absence of the additional VKEK and LKKK binding sites of the CAT domain.

Interestingly, we performed competition assays by analyzing the binding properties and activity of HSulf-2 in presence of isolated HD. Surprisingly, preliminary results showed that the addition of HD increased the binding of HSulf-2 to HS wish cells and its activity in our endosulfatase assay (data not shown). These data should be carefully confirmed, but could indicate the formation of HSulf-2/HD heterodimers, which would exhibit enhanced biological activities. Confirmation of these results would support further the ability of full length HSulf-

2 to dimerize, and prove that isolated, bacteria expressed HD would retain the structural features involved in the formation of HD/HD complexes.

Using CD, we compared the structural profile of HSulf2ΔSG, HSulf2ΔHD and HD-HSulf2. Results showed that HSulf2ΔHD was a totally folded protein, whereas HSulf2ΔSG featured unfolded region (Figure 51). We thus hypothesized that this region may correspond to the HD. It would be important to identify the disulfide bonds present in the isolated HD and compare them to those in HSulf-2 WT, in order to confirm the relevance of our produced HD.

The HD was divided in four constructs (HD-A, -B, -C, -D), to facilitate NMR study, and with the aim to identify the HP tetrasaccharide binding sites by 3D NMR. Results showed that all the constructs interacted with dp4 except for the HD-C (Figure 49). The assignment of HD-A and HD-D is in progress. For HD-B, immediate precipitation of the formed complex was observed, leading to the loss of the protein signal. To avoid this, we decided to perform the analysis using disaccharide (dp2) instead of tetrasaccharide, to reduce the binding affinity and decrease risks of precipitation. Preliminary results showed chemical shifts in the HD-B upon the addition of dp2 in a 1: 2.5 ratio (data not shown). This is in agreement with analysis of HD using the beads approach, which showed the interaction of three clusters with HP, present in HD-B and HD-D. It would be important in the future, to mutate these sites in HSulf-2 WT in order to understand their real implication.

5. HSulf-1 and HSulf-2

A final critical perspective of this work is to perform similar investigations on the HSulf-1 isoform, which I did not study extensively during my PhD. This was mostly due to lower expression and purification rates compared to the HSulf-2 isoform. Both enzymes are known to have similar enzyme activity *in vitro* but redundant or opposite functions *in vivo*. It is thus highly critical to clarify the molecular basis leading to such differences among isoforms. For this, several strategies could be proposed.

First, it would be important to characterize further the GAG/enzyme interactions and the substrate specificities for both isoforms. Many observations indicate that HSulf-1 and HSulf-2 do not target the same population of HS (Lamanna et al., 2006; Morimoto-Tomita et al., 2002). HSulf-1 tends to bind and remain on cell surface HS, whereas HSulf-2 is more rapidly released in the extracellular compartment and could thus act preferentially on ECM HS. One observation that may provide a rationale for this is that HSulf-1 has been reported to bind to non-substrate GAGs, which could sequester the enzyme (Figure 53, Milz et al., 2013). We failed to demonstrate this in our cellular assays. However, we could show that HSulf-1 and HSulf-2 did not bind in the same way to the cell surface of EA.hy926 endothelial cells. This clearly highlights the existence of difference substrate specificities towards HS substrates. To investigate this further, we could analyze the activity of HSulf-1 on the different size-defined oligosaccharides and the synthetic octasaccharides, and compare it to that of HSulf-2.

Second, VKEK and LKKK site of the CAT domain are well conserved among isoforms (see page 176). It would thus be interesting to study their implications in the endosulfatase activity of the enzyme.

Third, it would be critical to characterize the structural properties of HD-HSulf1, as performed for HD-HSulf2, since HD domain is the least conserved region of the proteins and is the most likely to govern isoform functional specificities. We have already produced HD-HSulf1, using our prokaryotic expression system and purification procedure. Preliminary data showed that HD-HSulf1 may not bind to oligosaccharides with the same affinity as HD-HSulf2.

Finally, contrary to HSulf-2, we demonstrated that HSulf-1 does not carry a GAG chain. Interestingly, HSulf-1 and HSulf-2 isoforms display opposite activities in cancer. Most studies have shown that HSulf-1 is downregulated in cancer and that an overexpression of

HSulf-1 reduced cancer cell proliferation of *in vitro* and *in vivo*. It would thus be important to assess the role of HSulf-1 in our *in vivo* model.

Finally, one of the long-term aims of the project would be to develop inhibitors with anti-tumoral activities specifically targeting the pro-oncogenic HSulf-2 isoform only. For this, it is thus critical to perform structural studies on HSulf-1 isoform, to help for the rational design of Sulf-2 specific inhibitors.

6. Collaborative projects

During my PhD, I had the opportunity to participate to three different projects, as part of collaborations, which aimed at analyzing the structure of HS in specific biological contexts. I therefore performed the extraction of HS chains from tissues, purification by weak anion exchange chromatography and determination of their disaccharide composition by RPIP-HPLC, using a protocol previously established by Dr. Romain Vivès. Briefly, the first project consisted on investigating the effect of high sodium intake in mice on the structure of renal HS (Hijmans et al., 2017). This work showed that sodium diet increased the sulfation of renal HS, converting them into pro-inflammatory mediators. The second project aimed at analyzing the structure of kidney HS following renal fibrosis (Ferreras et al., 2019). Results showed that HS 2-*O*-sulfation was increased whereas 3-*O*-sulfation was decreased during the development of renal fibrosis in a mouse model. These changes were attributed to modifications in the expression of HS sulfotransferase enzymes. We speculated that these sulfation profile modifications could affect EGF signaling, which is known to stimulate epithelial cell repair and motility. The third project was a collaborative work with L'Oréal Paris, which aimed at analyzing and comparing the CS and HS expression and structure from fine and thick hair follicles.

CHAPTER

VII

Appendices

1. Annexes

1.1. Sequence of HSulf-2

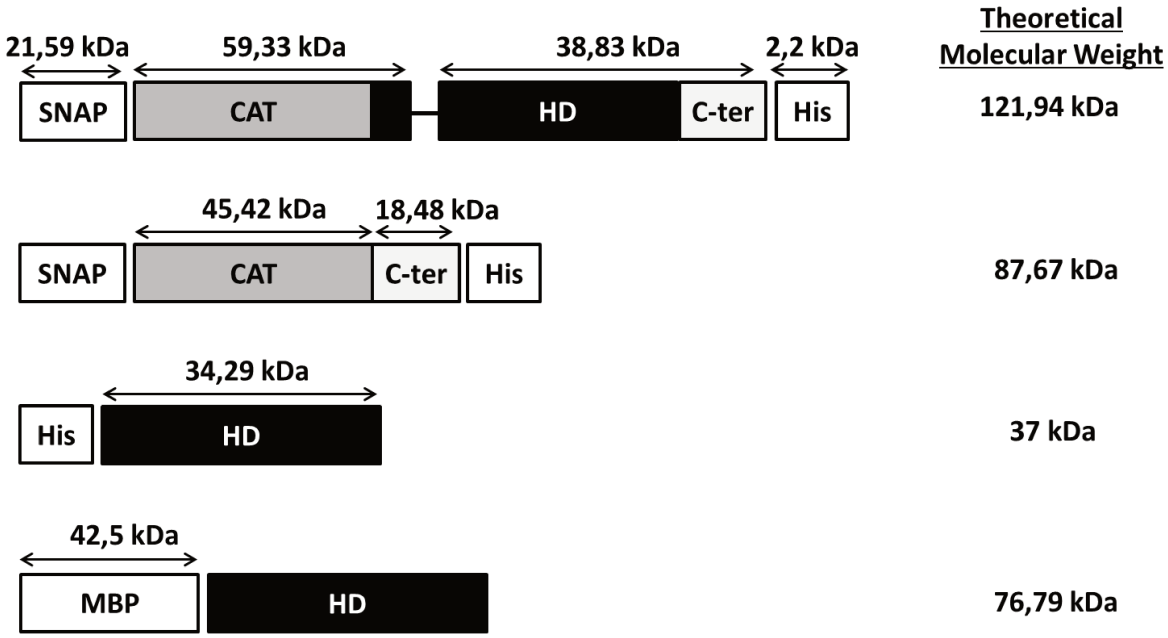
MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDD
QDVELGSMQVMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNT
YTNNENCSSPSWQAQHESRTFAVYL*N*STGYRTAFFGKYLNEYNGSYVPPGWKE
WVGLLKNSRFY*N*YTLCRNGVKEKHGSDYSKDYLTDLIT*N*DSVSFFRTSKKMYP
HRPVLMVISHAAPHGPEDSAPQYSRLFP*N*ASQHITPSYNYAPNPDKHWIMRYT
GPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNTYIVYTADH
GYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILD
IAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRD
NDKVDAQEENFLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKC
KGPMRLGGSRALSNLVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKAS
YVRSRSIR

SVAIEVDGRVYHVGLGDAAQPRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDY
SAANPIKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKN
LREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGLQEKDKVWL
LREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACT
SANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVL
NQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYEQYRQFQRRKWPEMKRP
SSKSLGQLWEGWEG

Signal peptide	Formylglycin converted cysteine
HP binding sites	*N-glycosylation*
O-glycosylation	**Disulfide bond**
Furin cleavage	HD domain

N terminal sequencing of remaining band after trypsin limited digestion

1.2. Different constructs of HSulf-2



1.3. Alignment of HSulf-1 and HSulf-2

```
HSulf-1   7    ALVLAVLGT---ELLGSLCSTVRSPRFRGRIQQERKNIRPNIILVLTDDQDVELGSLQVM  63
               +LVL +L      LLG   + +   R +GR Q++R+NIRPNIILVLTDDQDVELGS+QVM
HSulf-2   5    SLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQVM  64

HSulf-1   64   NKTRKIMEHGGATFINAFVTTPMCCPSRSSMLTGKYVHNHNVYTNNENCSSPSWQAMHEP  123
               NKTR+IME GGA FINAFVTTPMCCPSRSS+LTGKYVHNHN YTNNENCSSPSWQA HE
HSulf-2   65   NKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHES  124

HSulf-1   124  RTFAVYLNNTGYRTAFFGKYLNEYNGSYIPPGWREWLGLIKNSRFYNYTVCRNGIKEKHG  183
               RTFAVYLN+TGYRTAFFGKYLNEYNGSY+PPGW+EW+GL+KNSRFYNYT+CRNG+KEKHG
HSulf-2   125  RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHG  184

HSulf-1   184  FDYAKDYFTDLITNESINYFKMSKRMYPHRPVMMVISHAAPHGPEDSAPQFSKLYPNASQ  243
                DY+KDY TDLITN+S+++F+ SK+MYPHRPV+MVISHAAPHGPEDSAPQ+S+L+PNASQ
HSulf-2   185  SDYSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQ  244

HSulf-1   244  HITPSYNYAPNMDKHWIMQYTGPMLPIHMEFTNILQRKRLQTLMSVDDSVERLYNMLVET  303
               HITPSYNYAPN DKHWIM+YTGPM PIHMEFTN+LQRKRLQTLMSVDDS+E +YNMLVET
HSulf-2   245  HITPSYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVET  304

HSulf-1   304  GELENTYIIYTADHGYHIGQFGLVKGKSMPYDFDIRVPFFIRGPSVEPGSIVPQIVLNID  363
               GEL+NTYI+YTADHGYHIGQFGLVKGKSMPY+FDIRVPF++RGP+VE G + P IVLNID
HSulf-2   305  GELDNTYIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNID  364

HSulf-1   364  LAPTILDIAGLDTPPDVDGKSVLKLLDPEKPGNRFRTNKKAKIWRDTFLVERGKFLRKKE  423
               LAPTILDIAGLD P D+DGKS+LKLLD E+P NRF   KK ++WRD+FLVERGK L K++
HSulf-2   365  LAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRD  424

HSulf-1   424  ESSKNIQQSNHLPKYERVKELCQQARYQTACEQPGQKWQCIEDTSGKLRIHKCKGPSDLL  483
                   + Q+ N LPKY+RVK+LCQ+A YQTACEQ GQKWQC+ED +GKL++HKCKGP  L
HSulf-2   425  NDKVDAQEENFLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLG  484

HSulf-1   484  TVRQSTRNLYARGFHDKDKECSCRESGYRASRSQRKSQRQFLRNQGTPKYKPRFVHTRQT  543
                 R +  NL  + +    + C+C    Y+ S + R+ +    + + +      +V +R
HSulf-2   485  GSR-ALSNLVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFKRKYKAS------YVRSRSI  537

HSulf-1   544  RSLSVEFEGEIYDINLEEEEELQVLQPRNIAKRHDEGHKGPRDLQASSGGN-RGRMLADS  602
               RS+++E +G +Y + L +       QPRN+ KRH  G   P D     GG+  G
HSulf-2   538  RSVAIEVDGRVYHVGLGD-----AAQPRNLTKRHWPG--APEDQDDKDGGDFSGTGGLPD  590
```

```
HSulf-1  603  SNAVGPPTTVRVTHKCFILPNDSIHCERELYQSARAWKDHKAYIDKEIEALQDKIKNLRE  662
               +A  P   ++VTH+C+IL ND++ C+ +LY+S +AWKDHK +ID EIE LQ+KIKNLRE
HSulf-2  591  YSAANP---IKVTHRCYILENDTVQCDLDLYKSLQAWKDHKLHIDHEIETLQNKIKNLRE  647

HSulf-1  663  VRGHLKRRKPEECSCSKQSYYNKEKGVKKQEKLKSHLHPFKEAAQEVDSKLQLFKENNRR  722
              VRGHLK+++PEEC C K SY+ + KG  K     S LHPF++  QE D K+ L +E  R+
HSulf-2  648  VRGHLKKKRPEECDCHKISYHTQHKGRLKHR--GSSLHPFRKGLQEKD-KVWLLREQKRK  704

HSulf-1  723  RKKERKEKRRQRKGEECSLPGLTCFTHDNNHWQTAPFWNLGSFCACTSSNNNTYWCLRTV  782
              +K  +  KR Q   + CS+PGLTCFTHDN HWQTAPFW LG FCACTS+NNNTYWC+RT+
HSulf-2  705  KKLRRLLKRLQ-NNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFCACTSANNNTYWCMRTI  763

HSulf-1  783  NETHNFLFCEFATGFLEYFDMNTDPYQLTNTVHTVERGILNQLHVQLMELRSCQGYKQCN  842
              NETHNFLFCEFATGFLEYFD+NTDPYQL N V+T++R +LNQLHVQLMELRSC+GYKQCN
HSulf-2  764  NETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKGYKQCN  823

HSulf-1  843  PRPKNLDVGNKDGGSYDLHR------------------GQLWDGWEG  871
              PR +N+D+G KDGGSY+ +R                  GQLW+GWEG
HSulf-2  824  PRTRNMDLGLKDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG  870
```

Signal Peptide

Cysteine converted to Formylglycine

Heparin binding sites in CAT

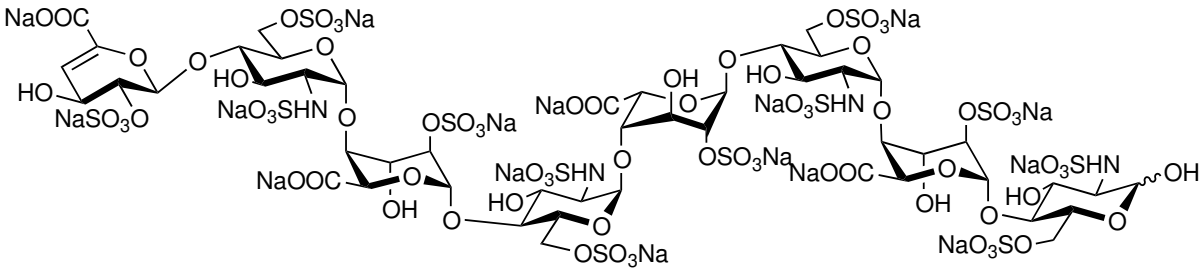
Heparin binding sites in HD

Furin cleavage

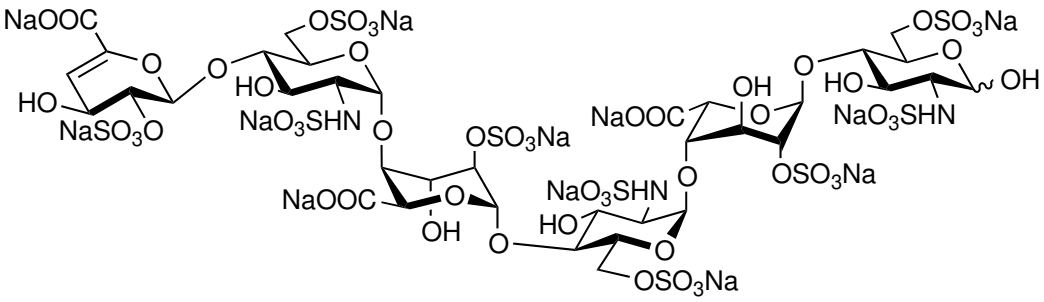
HD domain

1.4. Structure of natural oligosaccharides

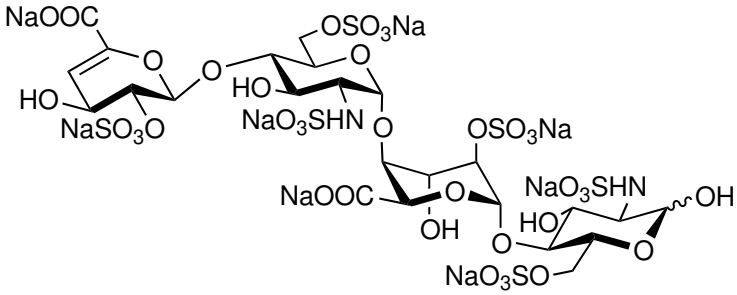
Octasaccharide (dp8)



Hexasaccharide (dp6)

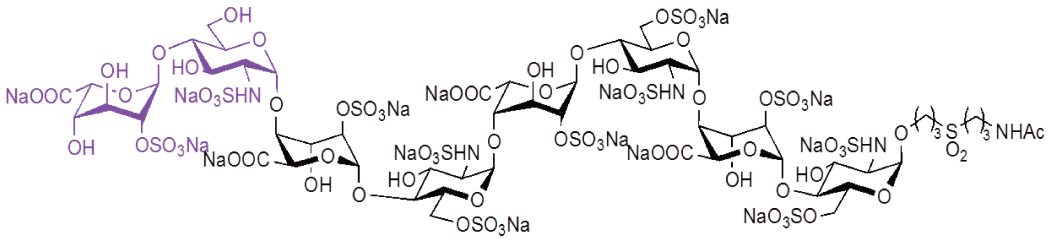


Tetrasaccharide (dp4)

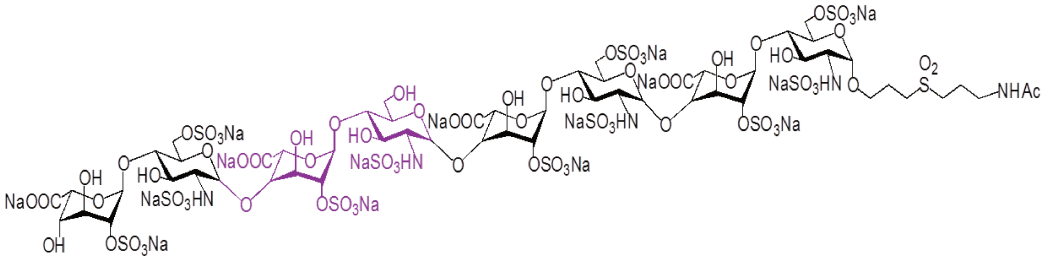


1.5. Structure of synthetic oligosaccharides

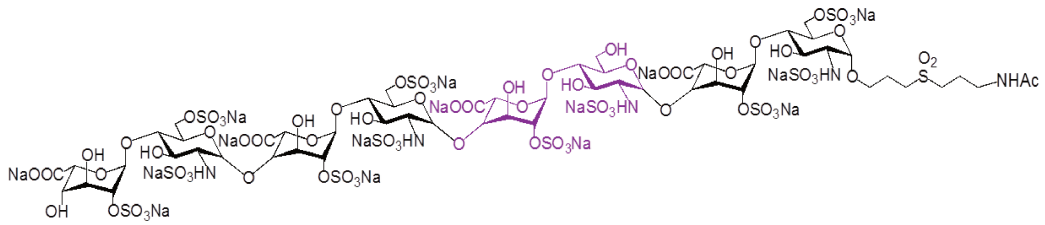
8.1



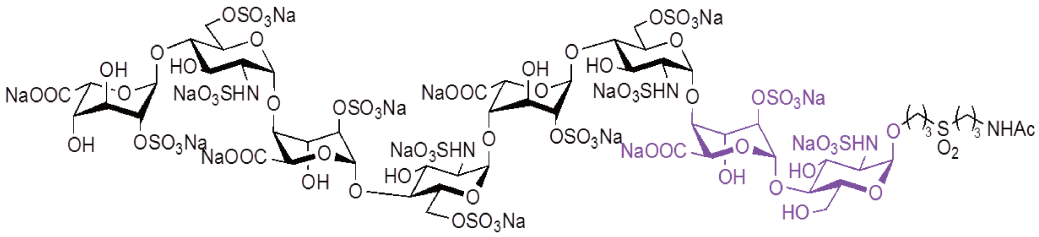
8.4



8.3



8.2



2. Résumé de la thèse en Français

2.1. Introduction

La vie des organismes pluricellulaires dépend largement des cellules qui communiquent entre elles et avec leur environnement. Cette interaction est médiée par un large éventail de protéines de signalisation telles que les facteurs de croissance, les interleukines, les cytokines, les chimiokines... qui se lient à leurs récepteurs apparentés afin de déclencher des voies de signalisation spécifiques. Pour avoir accès à leurs récepteurs, la plupart de ces protéines doivent traverser le glycocalyx, une épaisse couche de molécules glycosylées présentes à la surface de la cellule, et la matrice extracellulaire. Au sein de ces deux compartiments extracellulaires, les protéoglycanes (PGs) jouent un rôle central dans le contrôle de la diffusion et de l'activité de la plupart de ces protéines de signalisation. Les PGs sont formés d'un coeur protéique, auquel des polysaccharides anioniques linéaires qui appartiennent à la famille des glycosaminoglycanes (GAGs) sont fixés d'une manière covalente.

Les GAGs sont en effet caractérisés par leur capacité à interagir avec une grande variété de protéines de signalisation, contrôlant ainsi leur biodisponibilité, l'accès à leurs récepteurs, leur stockage, et leur protection contre la protéolyse. De ce fait, les PGs sont donc impliqués dans des processus cellulaires majeurs, tels que la prolifération cellulaire, la différenciation, migration, adhésion, chimioattraction, inflammation, réponses immunitaires, le contrôle de l'angiogenèse et de la coagulation. D'un point de vue structural, les GAGs sont constitués d'une répétition d'unités disaccharidiques composées d'une hexosamine et d'un acide uronique qui diffèrent selon le type de GAG. Les GAGs comprennent l'acide hyaluronique, la chondroïtine sulfate (CS), le dermatane sulfate (DS), le kératane sulfate, l'héparine (HP) et l'héparane sulfate (HS). Les HS représentent le GAG ayant les propriétés structurales et fonctionnelles les plus complexes. L'unité disaccharidique répétitive des HS est une glucosamine (GlcN) associée à un l'acide glucuronique (GlcA) ou un acide iduronique (IdoA), qui sont modifiés par l'addition de groupements sulfate dans différentes positions : en *N*- et 6-*O*- (plus rarement la sulfation 3-*O*-) des glucosamines et en 2-*O*- de l'acide uronique. Ces sulfatations définissent des régions saccharidiques spécifiques appelées domaines NS, définissant les sites de liaison pour les différents ligands protéiques du polysaccharide. Les HS sont donc constitués par une alternance de domaines non modifiés et de domaines NS. La structure des domaines NS est fortement régulée au cours de la biosynthèse des polysaccharides, un processus complexe qui a lieu dans le Golgi, et implique de nombreuses

enzymes. Ce processus génère une grande variété de structures d'HS et donc une large gamme de motifs de reconnaissance pour les différents ligands, permettant à la cellule de répondre finement aux stimuli de ses molécules de signalisation. Par exemple, la 6-*O*-sulfation des HS est cruciale pour la liaison et l'activation de nombreuses protéines, dont le VEGF, le FGF (facteur de croissance) ou le SDF (chimiokines).

La structure des HS est également contrôlée directement à la surface cellulaire par l'action d'une famille unique d'endosulfatases extracellulaires : les Sulfs. En altérant la structure des HS, ces enzymes modifient dramatiquement leurs propriétés d'interactions et sont de ce fait impliquées dans de nombreux processus physiopathologiques, notamment le cancer.

Les Sulfs existent sous la forme de deux isoformes chez l'homme : HSulf-1 et HSulf-2 qui sont caractérisées par une une organisation structurale très similaire. Les Sulfs, d'abord synthétisées sous forme de protéines de 125 kDa, deviennent matures après clivage par une protéase de type furine et comprennent deux sous-unités reliées entre elles par un ou plusieurs ponts disulfures. La sous unité N-terminale contient principalement le domaine catalytique (CAT) incluant le site actif de l'enzyme, et la sous unité C-terminale contient principalement un domaine hydrophile et basique (HD), responsable de la fixation des HS.

Cependant, malgré l'importance biologique de ces enzymes, la structure et le mode d'action des Sulfs demeurent énigmatiques. En particulier, dans le domaine du cancer, il a été montré que les deux formes de Sulfs humaines HSulf-1 et HSulf-2 dont l'activité enzymatique *in vitro* est très similaire, présentaient respectivement des propriétés anti- et pro-oncogénique *in vivo*. Le but de ce projet est donc de caractériser les propriétés structurales et fonctionnelles de l'enzyme HSulf-2 afin de mieux comprendre cet important système de régulation des HS.

2.2. Résultats et discussion

Récemment, nous avons mis en place un système d'expression et de purification de HSulf-2 sous forme recombinante. La production de l'enzyme est réalisée dans des cellules eucaryotes HEK293F en suspension, du fait de la présence des modifications post traductionnelles importantes pour l'acitivté de HSulf-2, telles que la cystéine convertie en résidu formylglycine (FGly), la coupure furine et les glycosylations. La culture en suspension dans un milieu dépourvu de sérum permet de faciliter la purification et d'augmenter les rendements d'expression. HSulf-2 est produite en présence d'une étiquette SNAP en N-terminal et d'un His tag en C-terminal. Malgré la présence de l'étiquette His, la purification

de la protéine n'a pas pu être réalisée par chromatographie d'affinité sur colonne de Nickel. Nous avons donc mis au point un protocole de purification en deux étapes : une selon la charge par chromatographie échangeuse d'ions, suivi par une autre étape selon la taille par chromatographie d'exclusion de taille. Grâce à cette approche, le rendement final de purification de HSulf-2 obtenu est d'environ 3 mg/L et le niveau de pureté de la protéine est de 95%. L'activité de l'enzyme est ensuite testée de trois manières différentes, en analysant : sa capacité à désulfater un substrat aryl, le 4MUS (activité arylsulfatase commune à toutes les sulfatases), sa capacité à digérer son substrat naturel (les HS, activité endosulfatase propre aux Sulfs) et enfin sa capacité à se lier aux HS.

a. Etude du mécanisme catalytique de HSulf-2

L'accès à une source de protéine recombinante nous a permis tout d'abord d'étudier les mécanismes de reconnaissance enzyme/substrat. Grâce à une technique de cartographie des sites d'interaction développée au laboratoire, nous avons identifié deux nouveaux motifs de reconnaissance des HS sur ces enzymes, dans le domaine CAT : V_{179}KEK et L_{401}KKK. Ces deux sites sont alignés avec le site actif (FGly) de l'enzyme. En plus de contenir le site actif de l'enzyme, le domaine CAT est donc capable de se lier aux HS. Des mutations ponctuelles de ces sites nous ont permis de confirmer leur contribution à l'activité enzymatique. Nous avons également produit de manière recombinante le domaine isolé CAT et nous avons étudié son activité. Un modèle du mode d'action de l'enzyme a été proposé à partir de ces données. Dans ce modèle, le HD se lie avec une forte affinité aux HS, et va les présenter au CAT, qui grâce à ses 2 sites VKEK et LKKK va guider la chaine vers le site catalytique afin de désulfater spécifiquement les groupements 6-*O*-S. Un article reprenant ces résultats a été publié en 2019 dans CMLS.

Nous avons ensuite voulu étudier la spécificité de substrat de HSulf-2. Pour cela, nous avons analysé la capacité de HSulf-2 ou de CAT à digérer des petits oligosaccharides qui diffèrent par leur taille (octasaccharides, hexasaccharides, tétrasaccharides et disaccharides). Concernant HSulf-2, l'enzyme a pu digérer tous les groupements 6-*O*-sulfates de tous les oligosaccharides, à l'exception du disaccharide. Ces résultats montrent qu'un tétrasaccharide correspond à la taille minimale de substrat pour l'activité de HSulf-2. La forme CAT a également pu reconnaitre tous les oligosaccharides, ce qui indique que le HD n'est pas indispensable à la reconnaissance des HS. Par contre, le CAT n'a pu catalyser l'élimination que d'un seul groupement sulfate sur le résidu terminal de chaque oligosaccharide. Ces

résultats suggèrent donc que le CAT seul présente une activité de type exosulfatase et que le HD est responsable de la processivité de l'enzyme le long de la chaine de HS.

b. Une nouvelle modification post-traductionnelle et son implication dans le cancer

L'accès à la protéine recombinante purifiée nous a permis d'identifier une modification post-traductionnelle uniquement présente chez HSulf-2 : la présence d'une chaine de CS, liée de manière covalente au HD. L'étape de purification par chromatographie d'exclusion de taille a en effet mis en évidence un poids moléculaire élevé inattendu pour l'enzyme (> ~ 1000 kDa, pour un poids moléculaire théorique de 98 kDa), qui ne pouvait pas être simplement attribué à des N-glycosylations. Une agrégation ou une multimérisation de la protéine ont été exclues. Il est important de noter qu'au cours de ces expériences, nous n'avions pas réussi à détecter la chaîne C-terminale contenant l'HD de l'enzyme par SDS-PAGE, bien que la présence des deux chaînes ait été vérifiée par séquençage N-terminal. Nous avons donc supposé que l'enzyme aurait pu être purifiée en complexe avec son substrat HS. Pour tester cela, HSulf-2 a été traitée avec des héparinases (pour digérer les substrats HS) ou de la chondroïtinase ABC (pour digérer des GAG non substrats de types CS / DS) avant la chromatographie d'exclusion de taille. Les résultats n'ont montré aucun effet du traitement par l'héparinase, alors que la digestion avec la chondroïtinase ABC réduisait considérablement le poids moléculaire de HSulf-2. Les tentatives de dissociation des complexes HSulf-2 / CS avec de l'urée ou de fortes concentrations de NaCl n'ont montré aucun effet, suggérant ainsi une liaison covalente entre le polysaccharide et la protéine. Par ailleurs, le traitement à la chondroïtinase permettait l'immunodétection de la chaîne C-terminale de l'enzyme par Western blot, localisant ainsi la présence de la chaîne sur le domaine HD. Cette modification pourrait jouer un rôle important dans la bio-distribution de l'enzyme dans les tissus. Pour confirmer cette hypothèse, nous avons produit un mutant de HSulf-2 dépourvu de cette chaine (HSulf2ΔSG) et étudié ses caractéristiques *in vitro* et *in vivo*. *In vitro*, les résultats ont montré que le mutant possédait une activité endosulfatase plus importante comparé à la forme sauvage (HSulf-2 WT). Nous avons donc émis l'hypothèse que l'activité accrue pourrait être due à des interactions de la chaîne de CS avec le HD, ou à une interférence électrostatique pouvant empêcher la liaison du substrat à ce domaine. Pour confirmer cela, nous avons comparé la liaison de HSulf-2 WT et HSulf2ΔSG à l'héparine par ELISA et aux HS de la surface cellulaire de cellules épithéliales Wish par FACS. Pour les deux approches, les résultats ont montré une augmentation significative de la liaison pour la

forme mutante HSulf2ΔSG. L'activité *in vitro* de HSulf-2 est donc affectée par sa chaîne de CS qui module les propriétés de reconnaissance enzyme / substrat.

Nous avons ensuite comparé les activités des formes HSulf-2 WT et HSulf2ΔSG sur la progression tumorale *in vivo*, en utilisant un modèle murin de xénogreffe orthotopique de tumeur mammaire. Pour cela, nous avons surexprimé HSulf-2 WT ou HSulf2ΔSG dans une lignée de cellules humaines de cancer du sein MDA-MB-231 ne produisant pas de HSulfs de manière endogène. Après sélection, les cellules transfectées avec les différentes formes de HSulf ont montré des niveaux d'expression similaires. L'activité des enzymes produites en MDA-MB-231 a été confirmée en analysant de l'héparine pré-incubée avec du milieu conditionné concentré provenant de cellules transfectées. Nous avons également confirmé la présence de la chaîne CS sur HSulf-2 WT. Des cellules MDA-MB-231 transfectées par HSulf-2 WT et mutante ont ensuite été injectées dans la glande mammaire de souris SCID, et l'apparition et la progression des tumeurs ont été suivies, jusqu'à ce que les tumeurs atteignent un volume de 1cm3. Les résultats ont montré peu d'effets de l'expression de HSulf-2 WT sur la taille de la tumeur. En revanche, les tumeurs exprimant le HSulf2ΔSG se sont développées plus rapidement et ont atteint une plus grande taille, par rapport aux tumeurs témoins et aux tumeurs HSulf-2 WT. Il convient de noter que les niveaux d'expression du HSulf-2 sont restés comparables dans les tumeurs WT et mutant. L'analyse histologique des sections tumorales à l'aide d'une coloration à l'éosine/hématoxine a montré une plus grande surface nécrosée dans les tumeurs témoins que dans les tumeurs exprimant HSulf. Comme la nécrose est une caractéristique de l'hypoxie dans les tumeurs en croissance, nous avons quantifié la vascularisation des tumeurs. Les résultats ont montré une augmentation significative de la vascularisation dans les tumeurs exprimants les HSulfs. Enfin, l'analyse de l'invasion tumorale dans le poumon, qui est une cible primaire pour les métastases dans ce modèle de tumeur, a montré que la taille des tumeurs secondaires induites par les métastases était significativement plus grande chez les tumeurs exprimant le mutant. Il est important de noter que de la protéolyse de HSulf-2 dans les tumeurs, conduisant à la perte de la chaine de CS a été observée dans les tumeurs exprimant HSulf-2 WT. L'ensemble de ces résultats montre que la chaîne CS nouvellement identifiée est une modification post traductionnelle fonctionnellement significative pour la régulation de HSulf-2 dans le cancer, qui atténue à la fois la croissance tumorale et l'invasion métastatique *in vivo*. Cependant, dans le microenvironnement tumoral, la chaîne CS peut être perdue par un traitement protéolytique par des métalloprotéases matricielles, conduisant à l'activation de HSulf-2, et augmentant la capacité des tumeurs à se développer et à former des métastases.

c. Etude structurale de HSulf-2

Un dernier aspect de mon travail de thèse a consisté à étudier la structure de HSulf-2. Au regard de la complexité des HSulfs, cette partie représentait un véritable défi scientifique et toutes les tentatives de cristallisation de ces enzymes se sont jusqu'à présent soldées par des échecs. Dans ce projet, nous avons choisi d'étudier séparément les domaines CAT et HD de l'enzyme. Le premier a été étudié par cristallographie aux rayons X, celui-ci étant fortement homologue à des sulfatases dont la structure a pu être résolue par cette approche. En parallèle, le domaine HD, dont les prédictions de séquences indiquent qu'il serait relativement peu structuré, a été analysé en RMN dans le but d'étudier son dynamisme.

Plusieurs essais de cristallisation ont été réalisés sur le domaine CAT jusqu'à présent. Nous avons essayé de cristalliser le CAT en présence d'une petite molécule très sulfaté nommée sulfocalixarene, qui contribue à stabiliser certaines protéines et donc à augmenter les chances d'obtenir des cristaux. Nous avons également essayé de produire le CAT dans des cellules HEK293S déficientes dans une glycosyl transferase, afin de limiter son niveau de glycosylation (CAT-S). Une réduction de glycosylation d'environ 6 kDa a ainsi été obtenue, tout en maintenant l'expression d'une enzyme active. Tous les essais de cristallisation réalisés sur le domaine CAT ont été sans succès. La stratégie de suppression du domaine HD n'a donc pas permis d'obtenir des cristaux diffractants. Le HD peut donc être important pour maintenir la structuration de la protéine entière. Nous avons donc décidé de réaliser une digestion ménagée de la protéine entière avec de la trypsine afin d'éliminer les régions flexibles exposées qui pourraient empêcher la cristallisation. Les résidus éliminés par la digestion à la trypsine sont les premiers résidus de HSulf-2, ainsi que le domaine HD. Ces résultats suggèrent ainsi que le HD peut être un domaine exposé flexible, sensible à la digestion de la trypsine. L'activité de l'enzyme n'a pas été affectée par la digestion. Ce produit de digestion a été purifié et soumis à la plateforme de cristallographie. Un cristal a pu être obtenu, mais les tests de diffraction réalisés au synchrotron Soleil ont délivré des spectres de diffraction typique du sel.

Concernant le HD, après un travail important, nous avons réussi à établir des conditions permettant l'expression et la purification de ce domaine en bactérie. Cette approche consiste à exprimer la protéine sous la forme de corps d'inclusion, et de réaliser un repliement *a posteriori*, par la technique de « flash refolding ». Des productions couplées à un marquage isotopique ont permis de réaliser les premières analyses structurales du domaine HD en RMN et de confirmer son caractère extrêmement dynamique. Il est très intéressant de noter que

nous avons observé la présence d'une région structurée pour des formes dégradées du HD, dégradation s'étant surtout produite dans la partie C-terminale du HD.

Nous avons ensuite ajouté le substrat du HD (des oligosaccharides d'héparine) dans le but de stabiliser les régions non structurées de ce domaine. Ces expériences ont été réalisées avec un tétrasaccharide d'héparine (dp4) ajouté au HD avec un rapport de 1:1. Les résultats ont montré que le degré de repliement de la protéine n'était pas modifié par l'ajout du substrat, mais nous avons observé certains déplacements chimiques, soulignant les interactions de résidus du HD avec l'oligosaccharide. Des travaux sont donc actuellement en cours afin d'identifier les sites de liaison du HD à l'héparine par RMN.

En parallèle, nous avons utilisé une technique de cartographie des sites d'interactions aux HS sur le HD. Les résultats ont montré la présence de trois clusters basiques dans le HD qui peuvent interagir avec les HS : R_{518}RKKKLFKKK, R_{649}GHLKKKKR et K_{702}RKKKKLRK. Sur la base de ces sites de liaison potentiels, nous avons conçu 4 constructions plus petites (~ 12 kDa, appelées HD-A, HD-B, HD-C et HD-D) chevauchantes et couvrant l'ensemble de la séquence du HD. Deux constructions contiennent les trois clusters identifiés (B et D), tandis que les deux autres ne les contiennent pas (A et C) (Figure 48). Nous avons produit et purifié les protéines HD A-D, de la même manière que la forme entière. Toutes les constructions ont donné, en RMN 2D, des spectres typiques de protéines non structurées, à l'exception de la HD-A, où certains signaux pouvant correspondre à des structures secondaires ont pu être détectés, comme cela l'avait déjà été observé pour la forme dégradée du HD. Nous avons ensuite analysé ces constructions en présence d'un tétrasaccharide et nous avons identifié celles pour lesquelles un déplacement chimique pouvait être observé, suite à une interaction avec le tétrasaccharide. Pour le HD-C, aucune variation n'a été observée, ce qui suggère fortement que le HD-C et le tétrasaccharide n'interagissent pas. Pour le HD-A et le HD-D, nous avons observé des décalages de certaines résonances, indiquant une interaction des deux espèces avec l'oligosaccharide. Pour le HD-B, l'addition du tétrasaccharide a entraîné une précipitation immédiate du complexe formé, mise en évidence par la perte complète du signal protéique. Nous avons ensuite entrepris une étude de RMN 3D pour identifier le site de liaison du dp4 sur les constructions HD-A et HD-D. HD-D a été entièrement assigné. Les résultats indiquent qu'à 10°C, le HD-D est essentiellement déplié, à l'exception de la partie C-terminale (E_{700}-R_{713}), qui montre une forte propension à former une structure hélicoïdale. Il est intéressant de noter que cette région contient le motif supposé de liaison à l'héparine. Ce travail est toujours en cours.

Le HD est produit dans un système procaryote, dépourvu de toutes les modifications post-traductionnelles. Habituellement, l'efficacité de repliement de la protéine repliée peut être estimée par son activité biologique, telle que l'activité enzymatique. Cependant, le HD n'a pas de site actif. Il était donc important de valider la pertinence structurale et fonctionnelle de ce domaine recombinant isolé. Sur le plan fonctionnel, nous avons comparé la capacité du HD et de Hsulf-2 à lier l'héparine par ELISA, et la fixation aux HS de la surface de cellules Wish par FACS. Les résultats obtenus ont montré que, comme HSulf-2, le HD pouvait se lier à l'héparine et aux HS, mais d'une manière moins efficace. Ces différences pourraient être dues à l'absence dans la forme HD des sites VKEK et LKKK du domaine CAT, qui sont des sites d'interaction supplémentaires pour la protéine entière, comme nous l'avons montré précédemment. Structuralement, nous avons effectué une analyse des formes HD, CAT et HSulf-2 par dichroïsme circulaire (CD) et par SAXS (pour les formes CAT et HSulf-2). Ces deux techniques nous ont permis d'évaluer les structures secondaires et les propriétés de repliement de ces protéines. Les résultats de CD ont montré que les profils de spectre de HSulf-2 et de CAT correspondaient à celui d'une protéine structurée contenant notamment des hélices α. Au contraire, le profil de spectre du HD indique une protéine désordonnée. Concernant les expériences de SAXS, l'analyse de CAT délivre un pic distinct en forme de cloche, indiquant une protéine globulaire. Pour HSulf-2, les résultats obtenus montrent un pic accompagné d'une remontée vers un plateau, suggérant une protéine partiellement globulaire avec une partie allongée. Ces résultats suggèrent donc que les régions dépliées trouvées dans HSulf-2 pourraient être principalement attribuées au domaine HD.

L'ensemble de ces travaux devrait nous permettre de mieux comprendre le mécanisme des désulfatation des HS par HSulf-2, l'effet pro-oncogénique de HSulf-2, et la structure de l'enzyme.

Acknowledgments

Tout d'abord, je remercie les membres de jury d'avoir accepté de prendre le temps pour évaluer mon travail de thèse : Pascale Marchot, Kenji Uchimura, Régis Daniel et Nicole Thielens.

Durant ma thèse, j'ai été chanceuse d'être entourée de gens formidables de l'IBS et d'ailleurs qui ont tous participé d'une manière ou d'une autre à l'achèvement de ce travail. Je ne vais surement pas pouvoir tous les citer mais je tiens à les remercier tous ! Je commence par remercier mon équipe SAGAG.

Mon directeur de thèse, Romain, tu m'as fait confiance dès le début et tu m'as donné toute l'autonomie dont j'avais besoin. En même temps, tu étais toujours présent pour suivre l'avancement du travail, pour lancer tes idées scientifiques super intéressantes et pour me donner tous les conseils nécessaires concernant les manips, la rédaction, et les présentations. Merci pour cet équilibre qui m'a permis de beaucoup apprendre. Merci également pour ton optimisme assez impressionnant et pour ton humour assez exceptionnel. Et je te le dis encore une fois « Tu es très drôle Romain !! ». Merci aussi pour ton grand cœur, c'est ce qui a compté le plus pour moi! Je savais que je pouvais tout te dire et tu prenais toujours le temps pour m'écouter, m'encourager et me rassurer! Je te dois aussi un dernier remerciement d'avoir été mon chef « hors IBS », pour me faire faire un baptême de plongée, de via ferrata et de ski (le ski était un peu moins réussi je dirais, mais ça me fera toujours trop rire !).

Je te remercie Hugues pour ta bonté, les remarques et les conseils pertinents que tu m'as donnés durant mes présentations, pour les sorties du labo inoubliables, et aussi pour m'avoir soutenue pour aller au congrès du GRC.

Rabia, merci pour ta gentillesse et ta bonne humeur, pour les discussions super agréables, pour tout ton aide et tes conseils, pendant ma thèse et aussi durant mon stage de M2. Merci aussi pour toute la nourriture extra délicieuse que t'as partagée avec nous ! Je trouve que tu es une vraie artiste avec plein de talents !

Evelyne, je suis très contente que tu sois devenue « SAGAG » et que nous soyons collègues de bureau. J'ai adoré pouvoir souvent discuter avec toi du travail et de plein d'autres choses. Merci pour ta gentillesse et ta bienveillance, ta spontanéité, ton aide dans les manips et toutes les astuces que tu m'as transmises !

Yoann, j'ai appris un autre type d'humour grâce à toi :P, merci pour toutes tes propositions d'aide ! Je te souhaite bon courage pour ta thèse !

Gaël, toujours souriant ! Merci pour tous les tampons que tu m'as préparés. Je te souhaite bon courage pour tes études.

Je tiens à remercier les anciens « SAGAG » :

Amal, je n'oublierai jamais ta disponibilité permanente, tes conseils et les discussions intéressantes sur Sulfs surtout pendant mon stage de M2. Merci également pour les moments très agréables !

Yoan Monneau, merci beaucoup d'avoir pris le temps de discuter avec moi de la purification du HD, de m'avoir appris plein de choses intéressantes et pour toutes les manips de RMN (et je n'oublie pas le tour sympa qu'on a fait à Paris !).

Merci aux stagiaires que j'ai encadrées, Elisa et Zahra. C'était un plaisir de faire cette expérience avec vous ! Une petite pensée également à: Damien, Élodie, Yanje, Sarah, Nassiba, Marie belle, Marina, Mayssa et Rayan.

Je trouve qu'avoir une bonne ambiance d'équipe aide énormément pour la motivation personnelle et professionnelle, et moi j'ai eu de la chance d'avoir fait ma thèse avec vous le SAGAG. Vous étiez une deuxième famille pour moi et vous allez beaucoup me manquez!

Je remercie ensuite les hbbs pour toutes les pauses de 17h et toutes les sorties. Vous êtes les meilleurs !!

Ma Lynda, tu m'as beaucoup manquée depuis que tu es partie! Je suis très chanceuse d'avoir fait ma thèse en même temps que toi! Je ne pouvais pas imaginer mieux. Merci pour tous les adorables moments qu'on a passés dans tous les coins de l'IBS et dans tous les pays visités, la nourriture et les gourmandises qu'on a partagées, les discussions interminables et les rires infinis sur tout et n'importe quoi (il n'y avait que nous qui comprenions nos blagues, pourtant elles étaient très drôles :P) et surtout merci pour ton amitié assez spéciale !

Rida, je suis très heureuse que t'aies rejoint le SAGAG pour ces 4 derniers mois. Merci ma chère pour la belle compagnie, de me faire beaucoup rire, de te rendre toujours disponible pour aider, et surtout pour nos promenades excitantes à carrefour après l'IBS :P

Kevin, mashallah l'énergie et l'humour :P merci de me faire peur/crier dans les couloirs de l'IBS, malheureusement je n'ai pas réussi à me venger!! Plus sérieusement, merci de m'écouter me plaindre quand ça ne va pas, de m'encourager à donner le meilleur de moi-même, de m'accompagner le soir quand je termine tard et de m'aider pour la mise en page de ce manuscrit !

Guillaume, cœur cœur :P, merci pour venir toujours nous voir, me donner des sourires avec ton humour et tes compliments libanais, et pour me sauver quand j'ai besoin de certains produits le soir. Kevin et guigui, votre version de « shou helou habibi shou helou» sera toujours la meilleure !

Rime, merci d'avoir toujours le sourire et l'énergie, tu me donnais des ondes positives à chaque fois que je te voyais ! Merci ma chère pour tous les moments très agréables, ta gentillesse et ton soutien !

Merci à tous les amis de l'IBS. Pleins de bons souvenirs avec chacun d'entre vous: Muge, Catarina, Stefaniia, Ilhem, Tomas, Momo, Mariam, Hind, Laura, Simon, Silvia, Anna, Justine,

Ricarda, Aldo, Wiktor, Sima, Sami, Fabien, Quentin, Marko, Orso, Audrey, Roberto, Julie, Anna, Kaiyao, Anne Sophie, Sebastian.

Merci également aux voisins de l'étage IRPAS, VIC et FDP, mais aussi aux RMN pour leur gentillesse et pour le partage du matériel de leur labo.

Je tiens à remercier tous mes collaborateurs, qui m'ont beaucoup aidé et avec qui j'ai eu le plaisir de travailler. Merci beaucoup, Odile, pour tout le temps que tu as consacré sur le projet *in vivo*, ta gentillesse et ton enthousiasme. Merci également, Caroline, pour ton aide avec les manips d'IHC, on a vite avancé grâce à toi! Merci beaucoup, Pierre, pour toutes les manips de RMN, ton efficacité, ta gentillesse et tes conseils pour les manips et la rédaction de la partie RMN du manuscrit. Merci également à Jérôme pour ses remarques importantes lors des réunions RMN. Merci, Régis, pour les analyses de SM, pour ta gentillesse et les conversations intéressantes, merci également à Ilham et Mélanie. Merci, Lionel, pour les essais de cell free, pour ta gentillesse et ton humour. Merci, Julien, pour les manips de CD et SAXS, pour te rendre disponible et pour ta motivation. Merci Françoise et Rose Laure pour la formation de M4D et Jean Philippe d'être toujours présent pour répondre à mes questions. Merci Jean-Pierre Andrieu pour les manips de séquençage N-terminal, Christine Le Narvor et David Bonaffé pour les oligosaccharides synthétiques, Luca Signor pour les manips de SM, Caroline Mas pour le MALLS, Michel Thépaut pour le DSF, Daphna Fenel pour la microscopie électronique, Alicia Vallet et Adrien Favier pour l'aide en manip RMN et Jean Baptise pour la formation SPR. Merci aux deux membres de mes CSI: Franck Fieschi et William Helbert pour le suivi de ma thèse, pour votre temps et vos conseils bien utiles.

Je termine mes remerciements par les gens hors IBS

Victoria, ma chère viki, merci de rester mon amie fidèle pendant ces années passées à Grenoble, d'être présente pour moi, toujours aussi serviable et aussi adorable.

Mes chers amis du Liban : Jamal et «أصدقاء الطفولة» : Lara, Maya, Mhamad, Mhamad et Wassim, merci de garder contact pendant toutes ces années et de me faire rire et oublier un peu le travail !

Un remerciement spécial à ma famille : mon père, ma mère, ma sœur et ma meilleure amie Reem, Mhamad, et ma grand-mère. Merci pour votre amour inconditionnel, votre support continu, vos prières sincères, merci pour les messages et les appels de tous les jours. Malgré la distance, je me suis toujours senti parmi vous, et ça m'a beaucoup aidé à avancer !

Mes deux petites nièces Sofia et Maria, merci de rendre la vie plus belle !

الله يخليلي ياكم عيلتي الحلوة .. بحبكم كتير

Bibliography

Abiatari, I., Kleeff, J., Li, J., Felix, K., Büchler, M.W., and Friess, H. (2006). Hsulf-1 regulates growth and invasion of pancreatic cancer cells. J. Clin. Pathol. *59*, 1052–1058.

Adhikari, N., Rusch, M., Mariash, A., Li, Q., Selleck, S.B., and Hall, J.L. (2008). Alterations in Heparan Sulfate in the Vessel in Response to Vascular Injury in the Mouse. J Cardiovasc. Transl. Res. *1*, 236–240.

Ai, X., Do, A.-T., Lozynska, O., Kusche-Gullberg, M., Lindahl, U., and Emerson, C.P. (2003). QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. J Cell Biol *162*, 341–351.

Ai, X., Do, A.-T., Kusche-Gullberg, M., Lindahl, U., Lu, K., and Emerson, C.P. (2006). Substrate Specificity and Domain Functions of Extracellular Heparan Sulfate 6-O-Endosulfatases, QSulf1 and QSulf2. J. Biol. Chem. *281*, 4969–4976.

Ai, X., Kitazawa, T., Do, A.-T., Kusche-Gullberg, M., Labosky, P.A., and Emerson, C.P. (2007). SULF1 and SULF2 regulate heparan sulfate-mediated GDNF signaling for esophageal innervation. Development *134*, 3327–3338.

Aikawa, J., and Esko, J.D. (1999). Molecular Cloning and Expression of a Third Member of the Heparan Sulfate/Heparin GlcNAcN-Deacetylase/ N-Sulfotransferase Family. J. Biol. Chem. *274*, 2690–2695.

Aikawa, J., Grobe, K., Tsujimoto, M., and Esko, J.D. (2001). Multiple Isozymes of Heparan Sulfate/Heparin GlcNAcN-Deacetylase/GlcN N-Sulfotransferase STRUCTURE AND ACTIVITY OF THE FOURTH MEMBER, NDST4. J. Biol. Chem. *276*, 5876–5882.

Alex, J.M., Rennie, M.L., Engilberge, S., Lehoczki, G., Dorottya, H., Fizil, Á., Batta, G., and Crowley, P.B. (2019). Calixarene-mediated assembly of a small antifungal protein. IUCrJ *6*, 238–247.

Alexopoulou, A.N., Multhaupt, H.A.B., and Couchman, J.R. (2007). Syndecans in wound healing, inflammation and vascular biology. Int. J. Biochem. Cell Biol. *39*, 505–528.

Alhasan, A.A., Spielhofer, J., Kusche-Gullberg, M., Kirby, J.A., and Ali, S. (2014). Role of 6-O-Sulfated Heparan Sulfate in Chronic Renal Fibrosis. J. Biol. Chem. *289*, 20295–20306.

Allen, B.L., and Rapraeger, A.C. (2003). Spatial and temporal expression of heparan sulfate in mouse development regulates FGF and FGF receptor assembly. J. Cell Biol. *163*, 637–648.

Allison, D.D., and Grande-Allen, K.J. (2006). Review. Hyaluronan: a powerful tissue engineering tool. Tissue Eng. *12*, 2131–2140.

Almond, A. (2007). Hyaluronan. Cell. Mol. Life Sci. *64*, 1591–1596.

Ambasta, R.K., Ai, X., and Emerson, C.P. (2007). Quail Sulf1 Function Requires Asparagine-linked Glycosylation. J. Biol. Chem. *282*, 34492–34499.

Ashikari-Hada, S., Habuchi, H., Kariya, Y., Itoh, N., Reddi, A.H., and Kimata, K. (2004). Characterization of Growth Factor-binding Structures in Heparin/Heparan Sulfate Using an Octasaccharide Library. J. Biol. Chem. *279*, 12346–12354.

Axelsson, J., Xu, D., Na Kang, B., Nussbacher, J.K., Handel, T.M., Ley, K., Sriramarao, P., and Esko, J.D. (2012). Inactivation of heparan sulfate 2-O-sulfotransferase accentuates neutrophil infiltration during acute inflammation in mice. Blood *120*, 1742–1751.

Bai, X., Bame, K.J., Habuchi, H., Kimata, K., and Esko, J.D. (1997). Turnover of Heparan Sulfate Depends on 2-O-Sulfation of Uronic Acids. J. Biol. Chem. *272*, 23172–23179.

Bame, K.J., Lidholt, K., Lindahl, U., and Esko, J.D. (1991a). Biosynthesis of heparan sulfate. Coordination of polymer-modification reactions in a Chinese hamster ovary cell mutant defective in N-sulfotransferase. J. Biol. Chem. *266*, 10287–10293.

Bame, K.J., Reddy, R.V., and Esko, J.D. (1991b). Coupling of N-deacetylation and N-sulfation in a Chinese hamster ovary cell mutant defective in heparan sulfate N-sulfotransferase. J. Biol. Chem. *266*, 12461–12468.

Bame, K.J., Zhang, L., David, G., and Esko, J.D. (1994). Sulphated and undersulphated heparan sulphate proteoglycans in a Chinese hamster ovary cell mutant defective in N-sulphotransferase. Biochem. J. *303*, 81–87.

Bartlett, A.H., and Park, P.W. (2010). Proteoglycans in host–pathogen interactions: molecular mechanisms and therapeutic implications. Expert Rev. Mol. Med. *12*, e5.

Berninsone, P., and Hirschberg, C.B. (1998). Heparan Sulfate/HeparinN-Deacetylase/N-Sulfotransferase THE N-SULFOTRANSFERASE ACTIVITY DOMAIN IS AT THE CARBOXYL HALF OF THE HOLOENZYME. J. Biol. Chem. *273*, 25556–25559.

Bethea, H.N., Xu, D., Liu, J., and Pedersen, L.C. (2008). Redirecting the substrate specificity of heparan sulfate 2-O-sulfotransferase by structurally guided mutagenesis. Proc. Natl. Acad. Sci. U. S. A. *105*, 18724–18729.

Bielicki, J., and Hopwood, J.J. (1991). Human liver N-acetylgalactosamine 6-sulphatase. Purification and characterization. Biochem. J. *279 (Pt 2)*, 515–520.

Biter, A.B., de la Peña, A.H., Thapar, R., Lin, J.Z., and Phillips, K.J. (2016). DSF Guided Refolding As A Novel Method Of Protein Production. Sci. Rep. *6*, 18906.

Blackhall, F.H., Merry, C.L.R., Lyon, M., Jayson, G.C., Folkman, J., Javaherian, K., and Gallagher, J.T. (2003). Binding of endostatin to endothelial heparan sulphate shows a differential requirement for specific sulphates. Biochem. J. *375*, 131–139.

Braga, T., Grujic, M., Lukinius, A., Hellman, L., Åbrink, M., and Pejler, G. (2007). Serglycin proteoglycan is required for secretory granule integrity in mucosal mast cells. Biochem. J. *403*, 49–57.

Bredrup, C., Knappskog, P.M., Majewski, J., Rødahl, E., and Boman, H. (2005). Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. Invest. Ophthalmol. Vis. Sci. *46*, 420–426.

Brickman, Y.G., Ford, M.D., Gallagher, J.T., Nurcombe, V., Bartlett, P.F., and Turnbull, J.E. (1998). Structural Modification of Fibroblast Growth Factor-binding Heparan Sulfate at a Determinative Stage of Neural Development. J. Biol. Chem. *273*, 4350–4359.

Bruinsma, I.B., te Riet, L., Gevers, T., ten Dam, G.B., van Kuppevelt, T.H., David, G., Küsters, B., de Waal, R.M.W., and Verbeek, M.M. (2010). Sulfation of heparan sulfate associated with amyloid-beta plaques in patients with Alzheimer's disease. Acta Neuropathol. (Berl.) *119*, 211–220.

Buczek-Thomas, J.A., Hsia, E., Rich, C.B., Foster, J.A., and Nugent, M.A. (2008). Inhibition of histone acetyltransferase by glycosaminoglycans. J. Cell. Biochem. *105*, 108–120.

Bullock, S.L., Fletcher, J.M., Beddington, R.S.P., and Wilson, V.A. (1998). Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. Genes Dev. *12*, 1894–1906.

Bülow, H.E., and Hobert, O. (2004). Differential Sulfations and Epimerization Define Heparan Sulfate Specificity in Nervous System Development. Neuron *41*, 723–736.

Buraschi, S., Neill, T., and Iozzo, R.V. (2019). Decorin is a devouring proteoglycan: Remodeling of intracellular catabolism via autophagy and mitophagy. Matrix Biol. *75–76*, 260–270.

Busse, M., Feta, A., Presto, J., Wilén, M., Grønning, M., Kjellén, L., and Kusche-Gullberg, M. (2007). Contribution of EXT1, EXT2, and EXTL3 to Heparan Sulfate Chain Elongation. J. Biol. Chem. *282*, 32802–32810.

Busse-Wicher, M., Wicher, K.B., and Kusche-Gullberg, M. (2014). The extostosin family: Proteins with many functions. Matrix Biol. *35*, 25–33.

Cadwallader, A.B., and Yost, H.J. (2006). Combinatorial expression patterns of heparan sulfate sulfotransferases in zebrafish: I. The 3-O-sulfotransferase family. Dev. Dyn. *235*, 3423–3431.

Capila, I., and Linhardt, R.J. (2002). Heparin-protein interactions. Angew. Chem. Int. Ed Engl. *41*, 391–412.

Capurro, M.I., Xiang, Y.-Y., Lobe, C., and Filmus, J. (2005). Glypican-3 Promotes the Growth of Hepatocellular Carcinoma by Stimulating Canonical Wnt Signaling. Cancer Res. *65*, 6245–6254.

Capurro, M.I., Xu, P., Shi, W., Li, F., Jia, A., and Filmus, J. (2008). Glypican-3 inhibits Hedgehog signaling during development by competing with patched for Hedgehog binding. Dev. Cell *14*, 700–711.

Cardin, A.D., and Weintraub, H.J. (1989). Molecular modeling of protein-glycosaminoglycan interactions. Arterioscler. Dallas Tex *9*, 21–32.

Carey, D.J. (1997). Syndecans: multifunctional cell-surface co-receptors. Biochem. J. *327*, 1–16.

Carlsson, P., Presto, J., Spillmann, D., Lindahl, U., and Kjellén, L. (2008). Heparin/Heparan Sulfate Biosynthesis PROCESSIVE FORMATION OF N-SULFATED DOMAINS. J. Biol. Chem. *283*, 20008–20014.

Caterson, B., and Melrose, J. (2018). Keratan sulfate, a complex glycosaminoglycan with unique functional capability. Glycobiology *28*, 182–206.

Celie, J.W.A.M., Rutjes, N.W.P., Keuning, E.D., Soininen, R., Heljasvaara, R., Pihlajaniemi, T., Dräger, A.M., Zweegman, S., Kessler, F.L., Beelen, R.H.J., et al. (2007). Subendothelial heparan sulfate proteoglycans become major L-selectin and monocyte chemoattractant protein-1 ligands upon renal ischemia/reperfusion. Am. J. Pathol. *170*, 1865–1878.

Chang, M.-C., Chen, C.-A., Hsieh, C.-Y., Lee, C.-N., Su, Y.-N., Hu, Y.-H., and Cheng, W.-F. (2009). Mesothelin inhibits paclitaxel-induced apoptosis through the PI3K pathway. Biochem. J. *424*, 449–458.

Chau, B.N., Diaz, R.L., Saunders, M.A., Cheng, C., Chang, A.N., Warrener, P., Bradshaw, J., Linsley, P.S., and Cleary, M.A. (2009). Identification of SULF2 as a novel transcriptional target of p53 by use of integrated genomic analyses. Cancer Res. *69*, 1368–1374.

Chen, L., and Sanderson, R.D. (2009). Heparanase Regulates Levels of Syndecan-1 in the Nucleus. PLOS ONE *4*, e4947.

Chen, Z., Fan, J.-Q., Li, J., Li, Q.-S., Yan, Z., Jia, X.-K., Liu, W.-D., Wei, L.-J., Zhang, F.-Z., Gao, H., et al. (2009). Promoter hypermethylation correlates with the Hsulf-1 silencing in human breast and gastric cancer. Int. J. Cancer *124*, 739–744.

Cole, C.L., Rushton, G., Jayson, G.C., and Avizienyte, E. (2014). Ovarian Cancer Cell Heparan Sulfate 6-O-Sulfotransferases Regulate an Angiogenic Program Induced by Heparin-binding Epidermal Growth Factor (EGF)-like Growth Factor/EGF Receptor Signaling. J. Biol. Chem. *289*, 10488–10501.

Conrad, A.H., Zhang, Y., Tasheva, E.S., and Conrad, G.W. (2010). Proteomic analysis of potential keratan sulfate, chondroitin sulfate A, and hyaluronic acid molecular interactions. Invest. Ophthalmol. Vis. Sci. *51*, 4500–4515.

Cosma, M.P., Pepe, S., Annunziata, I., Newbold, R.F., Grompe, M., Parenti, G., and Ballabio, A. (2003). The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. Cell *113*, 445–456.

Coulson-Thomas, V.J. (2016). The role of heparan sulphate in development: the ectodermal story. Int. J. Exp. Pathol. *97*, 213–229.

Dai, Y., Yang, Y., MacLeod, V., Yue, X., Rapraeger, A.C., Shriver, Z., Venkataraman, G., Sasisekharan, R., and Sanderson, R.D. (2005). HSulf-1 and HSulf-2 Are Potent Inhibitors of Myeloma Tumor Growth in Vivo. J. Biol. Chem. *280*, 40066–40073.

Danesin, C., Agius, E., Escalas, N., Ai, X., Emerson, C., Cochard, P., and Soula, C. (2006). Ventral Neural Progenitors Switch toward an Oligodendroglial Fate in Response to Increased Sonic Hedgehog (Shh) Activity: Involvement of Sulfatase 1 in Modulating Shh Signaling in the Ventral Spinal Cord. J. Neurosci. *26*, 5037–5048.

Dang, X., Zhu, Q., Wang, L., Su, H., Lin, H., Zhou, N., Liang, T., Wang, Z., Huang, S., Ren, Q., et al. (2009). Macular corneal dystrophy in a Chinese family related with novel mutations of CHST6. Mol. Vis. *15*, 700–705.

Debarnot, C., Monneau, Y.R., Roig-Zamboni, V., Delauzun, V., Le Narvor, C., Richard, E., Hénault, J., Goulet, A., Fadel, F., Vivès, R.R., et al. (2019). Substrate binding mode and catalytic mechanism of human heparan sulfate d-glucuronyl C5 epimerase. Proc. Natl. Acad. Sci. U. S. A. *116*, 6760–6765.

Deepa, S.S., Carulli, D., Galtrey, C., Rhodes, K., Fukuda, J., Mikami, T., Sugahara, K., and Fawcett, J.W. (2006). Composition of Perineuronal Net Extracellular Matrix in Rat Brain A DIFFERENT DISACCHARIDE COMPOSITION FOR THE NET-ASSOCIATED PROTEOGLYCANS. J. Biol. Chem. *281*, 17789–17800.

Dhoot, G.K., Gustafsson, M.K., Ai, X., Sun, W., Standiford, D.M., and Emerson, C.P. (2001). Regulation of Wnt Signaling and Embryo Patterning by an Extracellular Sulfatase. Science *293*, 1663–1666.

Dick, G., Tan, C.L., Alves, J.N., Ehlert, E.M.E., Miller, G.M., Hsieh-Wilson, L.C., Sugahara, K., Oosterhof, A., van Kuppevelt, T.H., Verhaagen, J., et al. (2013). Semaphorin 3A binds to the perineuronal nets via chondroitin sulfate type E motifs in rodent brains. J. Biol. Chem. *288*, 27384–27395.

Dierks, T., Schmidt, B., and von Figura, K. (1997). Conversion of cysteine to formylglycine: a protein modification in the endoplasmic reticulum. Proc. Natl. Acad. Sci. U. S. A. *94*, 11963–11968.

Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M.A., and Figura, K. von (1998a). Posttranslational Formation of Formylglycine in Prokaryotic Sulfatases by Modification of Either Cysteine or Serine. J. Biol. Chem. *273*, 25560–25564.

Dierks, T., Lecca, M.R., Schmidt, B., and von Figura, K. (1998b). Conversion of cysteine to formylglycine in eukaryotic sulfatases occurs by a common mechanism in the endoplasmic reticulum. FEBS Lett. *423*, 61–65.

Dierks, T., Lecca, M.R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999). Sequence determinants directing conversion of cysteine to formylglycine in eukaryotic sulfatases. EMBO J. *18*, 2084–2091.

Dierks, T., Schmidt, B., Borissenko, L.V., Peng, J., Preusser, A., Mariappan, M., and von Figura, K. (2003). Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. Cell *113*, 435–444.

DiGabriele, A.D., Lax, I., Chen, D.I., Svahn, C.M., Jaye, M., Schlessinger, J., and Hendrickson, W.A. (1998). Structure of a heparin-linked biologically active dimer of fibroblast growth factor. Nature *393*, 812.

Djerbal, L., Lortat-Jacob, H., and Kwok, J. (2017). Chondroitin sulfates and their binding molecules in the central nervous system. Glycoconj. J. *34*, 363–376.

Do, A.-T., Smeds, E., Spillmann, D., and Kusche-Gullberg, M. (2006). Overexpression of Heparan Sulfate 6-O-Sulfotransferases in Human Embryonic Kidney 293 Cells Results in Increased N-Acetylglucosaminyl 6-O-Sulfation. J. Biol. Chem. *281*, 5348–5356.

Dou, W., Xu, Y., Pagadala, V., Pedersen, L.C., and Liu, J. (2015). Role of Deacetylase Activity of N-Deacetylase/N-Sulfotransferase 1 in Forming N-Sulfated Domain in Heparan Sulfate. J. Biol. Chem. *290*, 20427–20437.

Edavettal, S.C., Lee, K.A., Negishi, M., Linhardt, R.J., Liu, J., and Pedersen, L.C. (2004). Crystal Structure and Mutational Analysis of Heparan Sulfate 3-O-Sulfotransferase Isoform 1. J. Biol. Chem. *279*, 25789–25797.

El Masri, R., Seffouh, A., Lortat-Jacob, H., and Vivès, R.R. (2017). The “in and out” of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate. Glycoconj. J. *34*, 285–298.

Esko, J.D., and Lindahl, U. (2001). Molecular diversity of heparan sulfate. J. Clin. Invest. *108*, 169–173.

Esko, J.D., and Selleck, S.B. (2002). Order Out of Chaos: Assembly of Ligand Binding Sites in Heparan Sulfate. Annu. Rev. Biochem. *71*, 435–471.

Faham, S., and Hileman, R. e. (1996). Heparin structure and interactions with basic fibroblast growth factor. Science *271*, 1116.

Fan, G., Xiao, L., Cheng, L., Wang, X., Sun, B., and Hu, G. (2000). Targeted disruption of NDST-1 gene leads to pulmonary hypoplasia and neonatal respiratory distress in mice. FEBS Lett. *467*, 7–11.

Fedarko, N.S., and Conrad, H.E. (1986). A unique heparan sulfate in the nuclei of hepatocytes: structural changes with the growth state of the cells. J. Cell Biol. *102*, 587–599.

Ferreras, C., Rushton, G., Cole, C.L., Babur, M., Telfer, B.A., Kuppevelt, T.H. van, Gardiner, J.M., Williams, K.J., Jayson, G.C., and Avizienyte, E. (2012). Endothelial Heparan Sulfate 6-O-Sulfation Levels Regulate Angiogenic Responses of Endothelial Cells to Fibroblast Growth Factor 2 and Vascular Endothelial Growth Factor. J. Biol. Chem. *287*, 36132–36146.

Ferreras, L., Moles, A., Situmorang, G.R., El Masri, R., Wilson, I.L., Cooke, K., Thompson, E., Kusche-Gullberg, M., Vivès, R.R., Sheerin, N.S., et al. (2019). Heparan sulfate in chronic kidney diseases: Exploring the role of 3-O-sulfation. Biochim. Biophys. Acta Gen. Subj. *1863*, 839–848.

Fey, J., Balleininger, M., Borissenko, L.V., Schmidt, B., von Figura, K., and Dierks, T. (2001). Characterization of posttranslational formylglycine formation by luminal components of the endoplasmic reticulum. J. Biol. Chem. *276*, 47021–47028.

Feyerabend, T.B., Jin-Ping Li, Lindahl, U., and Rodewald, H.-R. (2006). Heparan sulfate C5-epimerase is essential for heparin biosynthesis in mast cells. Nat. Chem. Biol. *2*, 195–196.

Feyzi, E., Lustig, F., Fager, G., Spillmann, D., Lindahl, U., and Salmivirta, M. (1997). Characterization of heparin and heparan sulfate domains binding to the long splice variant of platelet-derived growth factor A chain. J. Biol. Chem. *272*, 5518–5524.

Feyzi, E., Saldeen, T., Larsson, E., Lindahl, U., and Salmivirta, M. (1998). Age-dependent Modulation of Heparan Sulfate Structure and Function. J. Biol. Chem. *273*, 13395–13398.

Filmus, J., and Capurro, M. (2014). The role of glypicans in Hedgehog signaling. Matrix Biol. J. Int. Soc. Matrix Biol. *35*, 248–252.

Forsberg, E., Pejler, G., Ringvall, M., Lunderius, C., Tomasini-Johansson, B., Kusche-Gullberg, M., Eriksson, I., Ledin, J., Hellman, L., and Kjellén, L. (1999). Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme. Nature *400*, 773–776.

Freeman, S.D., Keino-Masu, K., Masu, M., and Ladher, R.K. (2015). Expression of the heparan sulfate 6-O-endosulfatases, Sulf1 and Sulf2, in the avian and mammalian inner ear suggests a role for sulfation during inner ear development. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *244*, 168–180.

Frese, M.-A., Milz, F., Dick, M., Lamanna, W.C., and Dierks, T. (2009). Characterization of the human sulfatase Sulf1 and its high affinity heparin/heparan sulfate interaction domain. J. Biol. Chem. *284*, 28033–28044.

Funderburgh, J.L. (2002). Keratan sulfate biosynthesis. IUBMB Life *54*, 187–194.

Gallagher, J. (2015). Fell-Muir Lecture: Heparan sulphate and the art of cell regulation: a polymer chain conducts the protein orchestra. Int. J. Exp. Pathol. *96*, 203–231.

Gandhi, N.S., and Mancera, R.L. (2008). The structure of glycosaminoglycans and their interactions with proteins. Chem. Biol. Drug Des. *72*, 455–482.

Ghiselli, G., and Agrawal, A. (2005). The human D-glucuronyl C5-epimerase gene is transcriptionally activated through the β-catenin–TCF4 pathway. Biochem. J. *390*, 493–499.

Gialeli, C., Theocharis, A.D., and Karamanos, N.K. (2011). Roles of matrix metalloproteinases in cancer progression and their pharmacological targeting. FEBS J. *278*, 16–27.

Grigorieva, E.V. (2011). D-Glucuronyl C5-epimerase suppresses small-cell lung cancer cell proliferation in vitro and tumour growth in vivo. Br. J. Cancer 9.

Grigorieva, E., Eshchenko, T., Rykova, V.I., Chernakov, A., Zabarovsky, E., and Sidorov, S.V. (2008). Decreased expression of human D-glucuronyl C5-epimerase in breast cancer. Int. J. Cancer *122*, 1172–1176.

Gubbiotti, M.A., Vallet, S.D., Ricard-Blum, S., and Iozzo, R.V. (2016). Decorin interacting network: A comprehensive analysis of decorin-binding partners and their versatile functions. Matrix Biol. *55*, 7–21.

Habuchi, O. (2000). Diversity and functions of glycosaminoglycan sulfotransferases. Biochim. Biophys. Acta BBA - Gen. Subj. *1474*, 115–127.

Habuchi, H., Suzuki, S., Saito, T., Tamura, T., Harada, T., Yoshida, K., and Kimata, K. (1992). Structure of a heparan sulphate oligosaccharide that binds to basic fibroblast growth factor. Biochem. J. *285*, 805–813.

Habuchi, H., Habuchi, O., and Kimata, K. (1995). Purification and Characterization of Heparan Sulfate 6-Sulfotransferase from the Culture Medium of Chinese Hamster Ovary Cells. J. Biol. Chem. *270*, 4172–4179.

Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K., and Kimata, K. (2000). The Occurrence of Three Isoforms of Heparan Sulfate 6-O-Sulfotransferase Having Different Specificities for Hexuronic Acid Adjacent to the TargetedN-Sulfoglucosamine. J. Biol. Chem. *275*, 2859–2868.

Habuchi, H., Miyake, G., Nogami, K., Kuroiwa, A., Matsuda, Y., Kusche-Gullberg, M., Habuchi, O., Tanaka, M., and Kimata, K. (2003). Biosynthesis of heparan sulphate with diverse structures and functions: two alternatively spliced forms of human heparan sulphate 6-O-sulphotransferase-2 having different expression patterns and properties. Biochem. J. *371*, 131–142.

Habuchi, H., Nagai, N., Sugaya, N., Atsumi, F., Stevens, R.L., and Kimata, K. (2007). Mice Deficient in Heparan Sulfate 6-O-Sulfotransferase-1 Exhibit Defective Heparan Sulfate Biosynthesis, Abnormal Placentation, and Late Embryonic Lethality. J. Biol. Chem. *282*, 15578–15588.

Habuchi, O., Hirahara, Y., Uchimura, K., and Fukuta, M. (1996). Enzymatic sulfation of galactose residue of keratan sulfate by chondroitin 6-sulfotransferase. Glycobiology *6*, 51–57.

Hagner-Mcwhirter, A., Lindahl, U., and Li, J. p (2000). Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. Biochem. J. *347*, 69–75.

Hagner-McWhirter, Å., Li, J.-P., Oscarson, S., and Lindahl, U. (2004). Irreversible Glucuronyl C5-epimerization in the Biosynthesis of Heparan Sulfate. J. Biol. Chem. *279*, 14631–14638.

Hammond, E., Khurana, A., Shridhar, V., and Dredge, K. (2014). The Role of Heparanase and Sulfatases in the Modification of Heparan Sulfate Proteoglycans within the Tumor Microenvironment and Opportunities for Novel Cancer Therapeutics. Front. Oncol. *4*.

Han, C.H., Huang, Y.-J., Lu, K.H., Liu, Z., Mills, G.B., Wei, Q., and Wang, L.-E. (2011). Polymorphisms in the SULF1 gene are associated with early age of onset and survival of ovarian cancer. J. Exp. Clin. Cancer Res. *30*, 5.

Hanson, S.R., Best, M.D., and Wong, C.-H. (2004). Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. Angew. Chem. Int. Ed Engl. *43*, 5736–5763.

Harder, A., Möller, A.-K., Milz, F., Neuhaus, P., Walhorn, V., Dierks, T., and Anselmetti, D. (2015). Catch bond interaction between cell-surface sulfatase Sulf1 and glycosaminoglycans. Biophys. J. *108*, 1709–1717.

Harmer, N.J. (2006). Insights into the role of heparan sulphate in fibroblast growth factor signalling. Biochem. Soc. Trans. *34*, 442–445.

Hassing, H.C., Mooij, H., Guo, S., Monia, B.P., Chen, K., Kulik, W., Dallinga-Thie, G.M., Nieuwdorp, M., Stroes, E.S.G., and Williams, K.J. (2012). Inhibition of hepatic Sulf2 in vivo: a novel strategy to correct diabetic dyslipidemia. Hepatol. Baltim. Md *55*, 1746–1753.

Hayano, S., Kurosaka, H., Yanagita, T., Kalus, I., Milz, F., Ishihara, Y., Islam, M.N., Kawanabe, N., Saito, M., Kamioka, H., et al. (2012). Roles of heparan sulfate sulfation in dentinogenesis. J. Biol. Chem. *287*, 12217–12229.

Hayashida, K., Parks, W.C., and Park, P.W. (2009). Syndecan-1 shedding facilitates the resolution of neutrophilic inflammation by removing sequestered CXC chemokines. Blood *114*, 3033–3043.

He, X., Khurana, A., Roy, D., Kaufmann, S., and Shridhar, V. (2014). Loss of HSulf-1 expression enhances tumorigenicity by inhibiting Bim expression in ovarian cancer. Int. J. Cancer J. Int. Cancer *135*, 1783–1789.

He, Y.Q., Sutcliffe, E.L., Bunting, K.L., Li, J., Goodall, K.J., Poon, I.K.A., Hulett, M.D., Freeman, C., Zafar, A., McInnes, R.L., et al. (2012). The endoglycosidase heparanase enters the nucleus of T lymphocytes and modulates H3 methylation at actively transcribed genes via the interplay with key chromatin modifying enzymes. Transcription *3*, 130–145.

Henrotin, Y., Mathy, M., Sanchez, C., and Lambert, C. (2010). Chondroitin Sulfate in the Treatment of Osteoarthritis: From in Vitro Studies to Clinical Recommendations. Ther. Adv. Musculoskelet. Dis. *2*, 335–348.

Hijmans, R.S., Shrestha, P., Sarpong, K.A., Yazdani, S., El Masri, R., de Jong, W.H.A., Navis, G., Vivès, R.R., and van den Born, J. (2017). High sodium diet converts renal proteoglycans into pro-inflammatory mediators in rats. PloS One *12*, e0178940.

Hileman, R.E., Fromm, J.R., Weiler, J.M., and Linhardt, R.J. (1998). Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. BioEssays News Rev. Mol. Cell. Dev. Biol. *20*, 156–167.

Holmborn, K., Ledin, J., Smeds, E., Eriksson, I., Kusche-Gullberg, M., and Kjellén, L. (2004). Heparan Sulfate Synthesized by Mouse Embryonic Stem Cells Deficient in NDST1 and NDST2 Is 6-O-Sulfated but Contains No N-Sulfate Groups. J. Biol. Chem. *279*, 42355–42358.

Holmborn, K., Habicher, J., Kasza, Z., Eriksson, A.S., Filipek-Gorniok, B., Gopal, S., Couchman, J.R., Ahlberg, P.E., Wiweger, M., Spillmann, D., et al. (2012). On the Roles and Regulation of Chondroitin Sulfate and Heparan Sulfate in Zebrafish Pharyngeal Cartilage Morphogenesis. J. Biol. Chem. *287*, 33905–33916.

Holst, C.R., Bou-Reslan, H., Gore, B.B., Wong, K., Grant, D., Chalasani, S., Carano, R.A., Frantz, G.D., Tessier-Lavigne, M., Bolon, B., et al. (2007). Secreted sulfatases Sulf1 and Sulf2 have overlapping yet essential roles in mouse neonatal survival. PloS One *2*, e575.

Hosono-Fukao, T., Ohtake-Niimi, S., Hoshino, H., Britschgi, M., Akatsu, H., Hossain, M.M., Nishitsuji, K., van Kuppevelt, T.H., Kimata, K., Michikawa, M., et al. (2012). Heparan sulfate subdomains that are degraded by Sulf accumulate in cerebral amyloid ß plaques of Alzheimer's disease: evidence from mouse models and patients. Am. J. Pathol. *180*, 2056–2067.

Huber, S., Winterhalter, K.H., and Vaughan, L. (1988). Isolation and sequence analysis of the glycosaminoglycan attachment site of type IX collagen. J. Biol. Chem. *263*, 752–756.

Humphries, D.E., Wong, G.W., Friend, D.S., Gurish, M.F., Qiu, W.T., Huang, C., Sharpe, A.H., and Stevens, R.L. (1999). Heparin is essential for the storage of specific granule proteases in mast cells. Nature *400*, 769–772.

Huynh, M.B., Morin, C., Carpentier, G., Garcia-Filipe, S., Talhas-Perret, S., Barbier-Chassefière, V., Kuppevelt, T.H. van, Martelly, I., Albanese, P., and Papy-Garcia, D. (2012). Age-related Changes in

Rat Myocardium Involve Altered Capacities of Glycosaminoglycans to Potentiate Growth Factor Functions and Heparan Sulfate-altered Sulfation. J. Biol. Chem. *287*, 11363–11373.

Imberty, A., Lortat-Jacob, H., and Pérez, S. (2007). Structural view of glycosaminoglycan–protein interactions. Carbohydr. Res. *342*, 430–439.

Iozzo, R.V., and Murdoch, A.D. (1996). Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *10*, 598–614.

Iozzo, R.V., and Schaefer, L. (2015). Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. Matrix Biol. J. Int. Soc. Matrix Biol. *42*, 11–55.

Ishihara, M. (1994). Structural requirements in heparin for binding and activation of FGF-1 and FGF-4 are different from that for FGF-2. Glycobiology *4*, 817–824.

Islam, M., Gor, J., Perkins, S.J., Ishikawa, Y., Bächinger, H.P., and Hohenester, E. (2013). The Concave Face of Decorin Mediates Reversible Dimerization and Collagen Binding. J. Biol. Chem. *288*, 35526–35533.

Itoh, Y., and Nagase, H. (2002). Matrix metalloproteinases in cancer. Essays Biochem. *38*, 21–36.

Jacobsson, I., Lindahl, U., Jensen, J.W., Rodén, L., Prihar, H., and Feingold, D.S. (1984). Biosynthesis of heparin. Substrate specificity of heparosan N-sulfate D-glucuronosyl 5-epimerase. J. Biol. Chem. *259*, 1056–1063.

Jayson, G.C., Lyon, M., Paraskeva, C., Turnbull, J.E., Deakin, J.A., and Gallagher, J.T. (1998). Heparan Sulfate Undergoes Specific Structural Changes during the Progression from Human Colon Adenoma to Carcinoma in Vitro. J. Biol. Chem. *273*, 51–57.

Jayson, G.C., Hansen, S.U., Miller, G.J., Cole, C.L., Rushton, G., Avizienyte, E., and Gardiner, J.M. (2015). Synthetic heparan sulfate dodecasaccharides reveal single sulfation site interconverts CXCL8 and CXCL12 chemokine biology †Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cc05222j. Chem. Commun. Camb. Engl. *51*, 13846–13849.

Jemth, P., Kreuger, J., Kusche-Gullberg, M., Sturiale, L., Giménez-Gallego, G., and Lindahl, U. (2002). Biosynthetic Oligosaccharide Libraries for Identification of Protein-binding Heparan Sulfate Motifs EXPLORING THE STRUCTURAL DIVERSITY BY SCREENING FOR FIBROBLAST GROWTH FACTOR (FGF) 1 AND FGF2 BINDING. J. Biol. Chem. *277*, 30567–30573.

Jemth, P., Smeds, E., Do, A.-T., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2003). Oligosaccharide Library-based Assessment of Heparan Sulfate 6-O-Sulfotransferase Substrate Specificity. J. Biol. Chem. *278*, 24371–24376.

Ji, W., Yang, J., Wang, D., Cao, L., Tan, W., Qian, H., Sun, B., Qian, Q., Yin, Z., Wu, M., et al. (2011). hSulf-1 Gene Exhibits Anticancer Efficacy through Negatively Regulating VEGFR-2 Signaling in Human Cancers. PLoS ONE *6*, e23274.

Jiang, W., Ishino, Y., Hashimoto, H., Keino-Masu, K., Masu, M., Uchimura, K., Kadomatsu, K., Yoshimura, T., and Ikenaka, K. (2017). Sulfatase 2 Modulates Fate Change from Motor Neurons to Oligodendrocyte Precursor Cells through Coordinated Regulation of Shh Signaling with Sulfatase 1. Dev. Neurosci. *39*, 361–374.

Johansson, F.K., Göransson, H., and Westermark, B. (2005). Expression analysis of genes involved in brain tumor progression driven by retroviral insertional mutagenesis in mice. Oncogene *24*, 3896–3905.

Kalia, M., Chandra, V., Rahman, S.A., Sehgal, D., and Jameel, S. (2009). Heparan sulfate proteoglycans are required for cellular binding of the hepatitis E virus ORF2 capsid protein and for viral infection. J. Virol. *83*, 12714–12724.

Kalus, I., Salmen, B., Viebahn, C., von Figura, K., Schmitz, D., D'Hooge, R., and Dierks, T. (2009). Differential involvement of the extracellular 6-O-endosulfatases Sulf1 and Sulf2 in brain development and neuronal and behavioural plasticity. J. Cell. Mol. Med. *13*, 4505–4521.

Kalus, I., Rohn, S., Puvirajesinghe, T.M., Guimond, S.E., Eyckerman-Kölln, P.J., ten Dam, G., van Kuppevelt, T.H., Turnbull, J.E., and Dierks, T. (2015). Sulf1 and Sulf2 Differentially Modulate Heparan Sulfate Proteoglycan Sulfation during Postnatal Cerebellum Development: Evidence for Neuroprotective and Neurite Outgrowth Promoting Functions. PLoS ONE *10*.

Kameyama, H., Uchimura, K., Yamashita, T., Kuwabara, K., Mizuguchi, M., Hung, S.-C., Okuhira, K., Masuda, T., Kosugi, T., Ohgita, T., et al. (2019). The Accumulation of Heparan Sulfate S-Domains in Kidney Transthyretin Deposits Accelerates Fibril Formation and Promotes Cytotoxicity. Am. J. Pathol. *189*, 308–319.

Kamimura, K., Fujise, M., Villa, F., Izumi, S., Habuchi, H., Kimata, K., and Nakato, H. (2001). Drosophila Heparan Sulfate 6-O-Sulfotransferase (dHS6ST) Gene STRUCTURE, EXPRESSION, AND FUNCTION IN THE FORMATION OF THE TRACHEAL SYSTEM. J. Biol. Chem. *276*, 17014–17021.

Kamimura, K., Koyama, T., Habuchi, H., Ueda, R., Masu, M., Kimata, K., and Nakato, H. (2006). Specific and flexible roles of heparan sulfate modifications in Drosophila FGF signaling. J. Cell Biol. *174*, 773–778.

Katta, K., Imran, T., Busse-Wicher, M., Grønning, M., Czajkowski, S., and Kusche-Gullberg, M. (2015). Reduced Expression of EXTL2, a Member of the Exostosin (EXT) Family of Glycosyltransferases, in Human Embryonic Kidney 293 Cells Results in Longer Heparan Sulfate Chains. J. Biol. Chem. *290*, 13168–13177.

Khurana, A., Liu, P., Mellone, P., Lorenzon, L., Vincenzi, B., Datta, K., Yang, B., Linhardt, R.J., Lingle, W., Chien, J., et al. (2011). HSulf-1 Modulates FGF-2 and Hypoxia Mediated Migration and Invasion of Breast Cancer Cells. Cancer Res. *71*, 2152–2161.

Khurana, A., McKean, H., Kim, H., Kim, S.-H., mcguire, J., Roberts, L.R., Goetz, M.P., and Shridhar, V. (2012a). Silencing of HSulf-2 expression in MCF10DCIS.com cells attenuate ductal carcinoma in situ progression to invasive ductal carcinoma in vivo. Breast Cancer Res. BCR *14*, R43.

Khurana, A., Tun, H.W., Marlow, L., Copland, J.A., Dredge, K., and Shridhar, V. (2012b). Hypoxia negatively regulates heparan sulfatase 2 expression in renal cancer cell lines. Mol. Carcinog. *51*, 565–575.

Khurana, A., Jung-Beom, D., He, X., Kim, S.-H., Busby, R.C., Lorenzon, L., Villa, M., Baldi, A., Molina, J., Goetz, M.P., et al. (2013a). Matrix detachment and proteasomal inhibitors diminish Sulf-2 expression in breast cancer cell lines and mouse xenografts. Clin. Exp. Metastasis *30*, 407–415.

Khurana, A., Beleford, D., He, X., Chien, J., and Shridhar, V. (2013b). Role of heparan sulfatases in ovarian and breast cancer. Am. J. Cancer Res. *3*, 34–45.

Kinnunen, T., Huang, Z., Townsend, J., Gatdula, M.M., Brown, J.R., Esko, J.D., and Turnbull, J.E. (2005). Heparan 2-O-sulfotransferase, hst-2, is essential for normal cell migration in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. *102*, 1507–1512.

Kitagawa, H., Tsutsumi, K., Tone, Y., and Sugahara, K. (1997). Developmental Regulation of the Sulfation Profile of Chondroitin Sulfate Chains in the Chicken Embryo Brain. J. Biol. Chem. *272*, 31377–31381.

Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999). The Tumor Suppressor EXT-like Gene EXTL2 Encodes an α1, 4-N-Acetylhexosaminyltransferase That TransfersN-Acetylgalactosamine and N-Acetylglucosamine to the Common Glycosaminoglycan-Protein Linkage Region THE KEY ENZYME FOR THE CHAIN INITIATION OF HEPARAN SULFATE. J. Biol. Chem. *274*, 13933–13937.

Kitagawa, H., Izumikawa, T., Mizuguchi, S., Dejima, K., Nomura, K.H., Egusa, N., Taniguchi, F., Tamura, J., Gengyo-Ando, K., Mitani, S., et al. (2007). Expression of rib-1, a Caenorhabditis elegans Homolog of the Human Tumor Suppressor EXT Genes, Is Indispensable for Heparan Sulfate Synthesis and Embryonic Morphogenesis. J. Biol. Chem. *282*, 8533–8544.

Kleinschmit, A., Koyama, T., Dejima, K., Hayashi, Y., Kamimura, K., and Nakato, H. (2010). Drosophila heparan sulfate 6-O endosulfatase regulates Wingless morphogen gradient formation. Dev. Biol. *345*, 204–214.

Knaust, A., Schmidt, B., Dierks, T., von Bülow, R., and von Figura, K. (1998). Residues critical for formylglycine formation and/or catalytic activity of arylsulfatase A. Biochemistry *37*, 13941–13946.

Koike, T., Izumikawa, T., Sato, B., and Kitagawa, H. (2014). Identification of Phosphatase That Dephosphorylates Xylose in the Glycosaminoglycan-Protein Linkage Region of Proteoglycans. J. Biol. Chem. *289*, 6695–6708.

Kolset, S.O., and Gallagher, J.T. (1990). Proteoglycans in haemopoietic cells. Biochim. Biophys. Acta *1032*, 191–211.

Kolset, S.O., and Pejler, G. (2011). Serglycin: A Structural and Functional Chameleon with Wide Impact on Immune Cells. J. Immunol. *187*, 4927–4933.

Kovalszky, I., Dudás, J., Oláh-Nagy, J., Pogány, G., Töváry, J., Timár, J., Kopper, L., Jeney, A., and Iozzo, R.V. (1998). Inhibition of DNA topoisomerase I activity by heparan sulfate and modulation by basic fibroblast growth factor. Mol. Cell. Biochem. *183*, 11–23.

Krenn, E.C., Wille, I., Gesslbauer, B., Poteser, M., van Kuppevelt, T.H., and Kungl, A.J. (2008). Glycanogenomics: A qPCR-approach to investigate biological glycan function. Biochem. Biophys. Res. Commun. *375*, 297–302.

Kreuger, J., Salmivirta, M., Sturiale, L., Giménez-Gallego, G., and 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.

Kumagai, S., Ishibashi, K., Kataoka, M., Oguro, T., Kiko, Y., Yanagida, T., Aikawa, K., and Kojima, Y. (2016). Impact of Sulfatase-2 on cancer progression and prognosis in patients with renal cell carcinoma. Cancer Sci. *107*, 1632–1641.

Kusche, M., Bäckström, G., Riesenfeld, J., Petitou, M., Choay, J., and Lindahl, U. (1988). Biosynthesis of heparin. O-sulfation of the antithrombin-binding region. J. Biol. Chem. *263*, 15474–15484.

Kusche-Gullberg, M., and Kjellén, L. (2003). Sulfotransferases in glycosaminoglycan biosynthesis. Curr. Opin. Struct. Biol. *13*, 605–611.

Laabs, T., Carulli, D., Geller, H.M., and Fawcett, J.W. (2005). Chondroitin sulfate proteoglycans in neural development and regeneration. Curr. Opin. Neurobiol. *15*, 116–120.

Lai, J., Chien, J., Staub, J., Avula, R., Greene, E.L., Matthews, T.A., Smith, D.I., Kaufmann, S.H., Roberts, L.R., and Shridhar, V. (2003). Loss of HSulf-1 Up-regulates Heparin-binding Growth Factor Signaling in Cancer. J. Biol. Chem. *278*, 23107–23117.

Lai, J.-P., Chien, J.R., Moser, D.R., Staub, J.K., Aderca, I., Montoya, D.P., Matthews, T.A., Nagorney, D.M., Cunningham, J.M., Smith, D.I., et al. (2004a). hSulf1 Sulfatase promotes apoptosis of hepatocellular cancer cells by decreasing heparin-binding growth factor signaling. Gastroenterology *126*, 231–248.

Lai, J.-P., Chien, J., Strome, S.E., Staub, J., Montoya, D.P., Greene, E.L., Smith, D.I., Roberts, L.R., and Shridhar, V. (2004b). HSulf-1 modulates HGF-mediated tumor cell invasion and signaling in head and neck squamous carcinoma. Oncogene *23*, 1439–1447.

Lai, J.-P., Yu, C., Moser, C.D., Aderca, I., Han, T., Garvey, T.D., Murphy, L.M., Garrity-Park, M.M., Shridhar, V., Adjei, A.A., et al. (2006). SULF1 inhibits tumor growth and potentiates the effects of histone deacetylase inhibitors in hepatocellular carcinoma. Gastroenterology *130*, 2130–2144.

Lai, J.-P., Sandhu, D.S., Yu, C., Han, T., Moser, C.D., Jackson, K.K., Guerrero, R.B., Aderca, I., Isomoto, H., Garrity-Park, M.M., et al. (2008). Sulfatase 2 Up-Regulates Glypican 3, Promotes Fibroblast Growth Factor Signaling, and Decreases Survival in Hepatocellular Carcinoma. Hepatol. Baltim. Md *47*, 1211–1222.

Lamanna, W.C., Baldwin, R.J., Padva, M., Kalus, I., Ten Dam, G., van Kuppevelt, T.H., Gallagher, J.T., von Figura, K., Dierks, T., and Merry, C.L.R. (2006). Heparan sulfate 6-O-endosulfatases: discrete in vivo activities and functional co-operativity. Biochem. J. *400*, 63–73.

Lamanna, W.C., Frese, M.-A., Balleininger, M., and Dierks, T. (2008). Sulf Loss Influences N-, 2-O-, and 6-O-Sulfation of Multiple Heparan Sulfate Proteoglycans and Modulates Fibroblast Growth Factor Signaling. J. Biol. Chem. *283*, 27724–27735.

Langsdorf, A., Schumacher, V., Shi, X., Tran, T., Zaia, J., Jain, S., Taglienti, M., Kreidberg, J.A., Fine, A., and Ai, X. (2011). Expression regulation and function of heparan sulfate 6-O-endosulfatases in the spermatogonial stem cell niche. Glycobiology *21*, 152–161.

Laurent, U.B.G., and Reed, R.K. (1991). Turnover of hyaluronan in the tissues. Adv. Drug Deliv. Rev. *7*, 237–256.

Ledin, J., Ringvall, M., Thuveson, M., Eriksson, I., Wilén, M., Kusche-Gullberg, M., Forsberg, E., and Kjellén, L. (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.

Lemire, J.M., Chan, C.K., Bressler, S., Miller, J., LeBaron, R.G., and Wight, T.N. (2007). Interleukin-1β selectively decreases the synthesis of versican by arterial smooth muscle cells. J. Cell. Biochem. *101*, 753–766.

Lemjabbar-Alaoui, H., van Zante, A., Singer, M.S., Xue, Q., Wang, Y.-Q., Tsay, D., He, B., Jablons, D.M., and Rosen, S.D. (2010). Sulf-2, a heparan sulfate endosulfatase, promotes human lung carcinogenesis. Oncogene *29*, 635–646.

Leonova, E.I., and Galzitskaia, O.V. (2013). [Comparative characteristics of the structure and function for syndecan-1 from animal organisms]. Mol. Biol. (Mosk.) *47*, 505–512.

Levy-Adam, F., Feld, S., Cohen-Kaplan, V., Shteingauz, A., Gross, M., Arvatz, G., Naroditsky, I., Ilan, N., Doweck, I., and Vlodavsky, I. (2010). Heparanase 2 Interacts with Heparan Sulfate with High Affinity and Inhibits Heparanase Activity. J. Biol. Chem. *285*, 28010–28019.

Li, J.-P., and Kusche-Gullberg, M. (2016). Chapter Six - Heparan Sulfate: Biosynthesis, Structure, and Function. In International Review of Cell and Molecular Biology, K.W. Jeon, ed. (Academic Press), pp. 215–273.

Li, J., Kleeff, J., Abiatari, I., Kayed, H., Giese, N.A., Felix, K., Giese, T., Büchler, M.W., and Friess, H. (2005). Enhanced levels of Hsulf-1 interfere with heparin-binding growth factor signaling in pancreatic cancer. Mol. Cancer *4*, 14.

Li, J., Mo, M.-L., Chen, Z., Yang, J., Li, Q.-S., Wang, D.-J., Zhang, H., Ye, Y.-J., Li, H.-L., Zhang, F., et al. (2011). HSulf-1 inhibits cell proliferation and invasion in human gastric cancer. Cancer Sci. *102*, 1815–1821.

Li, J.-P., Gong, F., Hagner-McWhirter, Å., Forsberg, E., Åbrink, M., Kisilevsky, R., Zhang, X., and Lindahl, U. (2003). Targeted Disruption of a Murine Glucuronyl C5-epimerase Gene Results in Heparan Sulfate Lacking l-Iduronic Acid and in Neonatal Lethality. J. Biol. Chem. *278*, 28363–28366.

Li, Y., Sun, C., Yates, E.A., Jiang, C., Wilkinson, M.C., and Fernig, D.G. (2016). Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification. Open Biol. *6*.

Lidholt, K., Eriksson, I., and Kjellén, L. (1995). Heparin proteoglycans synthesized by mouse mastocytoma contain chondroitin sulphate. Biochem. J. *311*, 233–238.

Lin, X., Wei, G., Shi, Z., Dryer, L., Esko, J.D., Wells, D.E., and Matzuk, M.M. (2000). Disruption of Gastrulation and Heparan Sulfate Biosynthesis in EXT1-Deficient Mice. Dev. Biol. *224*, 299–311.

Lindahl, U. (2014). A personal voyage through the proteoglycan field. Matrix Biol. *35*, 3–7.

Lindahl, U., and Li, J. (2009). Interactions between heparan sulfate and proteins-design and functional implications. Int. Rev. Cell Mol. Biol. *276*, 105–159.

Lindahl, B., Eriksson, L., and Lindahl, U. (1995). Structure of heparan sulphate from human brain, with special regard to Alzheimer's disease. Biochem. J. *306*, 177–184.

Lindahl, B., Westling, C., Giménez-Gallego, G., Lindahl, U., and Salmivirta, M. (1999). Common Binding Sites for β-Amyloid Fibrils and Fibroblast Growth Factor-2 in Heparan Sulfate from Human Cerebral Cortex. J. Biol. Chem. *274*, 30631–30635.

Lindahl, U., Jacobsson, I., Hőők, M., Backstrőm, G., and Feingold, D.S. (1976). Biosynthesis of heparin. Loss of C-5 hydrogen during conversion of d-glucuronic to l-iduronic acid residues. Biochem. Biophys. Res. Commun. *70*, 492–499.

Liu, C., Sheng, J., Krahn, J.M., Perera, L., Xu, Y., Hsieh, P.-H., Dou, W., Liu, J., and Pedersen, L.C. (2014). Molecular Mechanism of Substrate Specificity for Heparan Sulfate 2-O-Sulfotransferase. J. Biol. Chem. *289*, 13407–13418.

Liu, H., Fu, X., Ji, W., Liu, K., Bao, L., Yan, Y., Wu, M., Yang, J., and Su, C. (2013). Human sulfatase-1 inhibits the migration and proliferation of SMMC-7721 hepatocellular carcinoma cells by downregulating the growth factor signaling. Hepatol. Res. Off. J. Jpn. Soc. Hepatol. *43*, 516–525.

Liu, J., Shworak, N.W., Sinay, P., Schwartz, J.J., Zhang, L., Fritze, L.M.S., and Rosenberg, R.D. (1999). Expression of Heparan Sulfate d-Glucosaminyl 3-O-Sulfotransferase Isoforms Reveals Novel Substrate Specificities. J. Biol. Chem. *274*, 5185–5192.

Liu, P., Khurana, A., Rattan, R., He, X., Kalloger, S., Dowdy, S., Gilks, B., and Shridhar, V. (2009). Regulation of HSulf-1 Expression by Variant Hepatic Nuclear Factor 1 in Ovarian Cancer. Cancer Res. *69*, 4843–4850.

Lortat-Jacob, H., Grosdidier, A., and Imberty, A. (2002). Structural diversity of heparan sulfate binding domains in chemokines. Proc. Natl. Acad. Sci. U. S. A. *99*, 1229–1234.

Lu, J., Auduong, L., White, E.S., and Yue, X. (2014). Up-Regulation of Heparan Sulfate 6-O-Sulfation in Idiopathic Pulmonary Fibrosis. Am. J. Respir. Cell Mol. Biol. *50*, 106–114.

Lui, N.S., Zante, A. van, Rosen, S.D., Jablons, D.M., and Lemjabbar-Alaoui, H. (2012). SULF2 expression by immunohistochemistry and overall survival in oesophageal cancer: a cohort study. BMJ Open *2*, e001624.

Lum, D.H., Tan, J., Rosen, S.D., and Werb, Z. (2007). Gene Trap Disruption of the Mouse Heparan Sulfate 6-O-Endosulfatase Gene, Sulf2. Mol. Cell. Biol. *27*, 678–688.

Lundin, L., Larsson, H., Kreuger, J., Kanda, S., Lindahl, U., Salmivirta, M., and Claesson-Welsh, L. (2000). Selectively Desulfated Heparin Inhibits Fibroblast Growth Factor-induced Mitogenicity and Angiogenesis. J. Biol. Chem. *275*, 24653–24660.

Lyon, M., Deakin, J.A., Mizuno, K., Nakamura, T., and Gallagher, J.T. (1994). Interaction of hepatocyte growth factor with heparan sulfate. Elucidation of the major heparan sulfate structural determinants. J. Biol. Chem. *269*, 11216–11223.

Lyon, M., Rushton, G., and Gallagher, J.T. (1997). The Interaction of the Transforming Growth Factor-βs with Heparin/Heparan Sulfate Is Isoform-specific. J. Biol. Chem. *272*, 18000–18006.

Lyon, M., Deakin, J.A., Rahmoune, H., Fernig, D.G., Nakamura, T., and Gallagher, J.T. (1998). Hepatocyte Growth Factor/Scatter Factor Binds with High Affinity to Dermatan Sulfate. J. Biol. Chem. *273*, 271–278.

MA, H.-Y., ZHANG, F., LI, J., MO, M.-L., CHEN, Z., LIU, L., ZHOU, H.-M., and SHENG, Q. (2011). HSulf-1 suppresses cell growth and down-regulates Hedgehog signaling in human gastric cancer cells. Oncol. Lett. *2*, 1291–1295.

Maccarana, M., Casu, B., and Lindahl, U. (1993). Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor. J. Biol. Chem. *268*, 23898–23905.

Mahmoud, S., Ibrahim, M., Hago, A., Huang, Y., Wei, Y., Zhang, J., Zhang, Q., Xiao, Y., Wang, J., Adam, M., et al. (2016). Overexpression of sulfatase-1 in murine hepatocarcinoma Hca-F cell line downregulates mesothelin and leads to reduction in lymphatic metastasis, both in vitro and in vivo. Oncotarget *7*, 75052–75063.

Mahmoud, S.A., Ibrahim, M.M., Musa, A.H., Huang, Y., Zhang, J., Wang, J., Wei, Y., Wang, L., Zhou, S., Xin, B., et al. (2018). Sulfatase-1 knockdown promotes in vitro and in vivo aggressive behavior of murine hepatocarcinoma Hca-P cells through up-regulation of mesothelin. J. Cell Commun. Signal. *12*, 603–613.

Maltseva, I., Chan, M., Kalus, I., Dierks, T., and Rosen, S.D. (2013). The SULFs, extracellular sulfatases for heparan sulfate, promote the migration of corneal epithelial cells during wound repair. PloS One *8*, e69642.

Mascellani, G., Liverani, L., Bianchini, P., Parma, B., Torri, G., Bisio, A., Guerrini, M., and Casu, B. (1993). Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate. Biochem. J. *296*, 639–648.

Matalon, R., Arbogast, B., and Dorfman, A. (1974). Deficiency of chondroitin sulfate N-acetylgalactosamine 4-sulfate sulfatase in Maroteaux-Lamy syndrome. Biochem. Biophys. Res. Commun. *61*, 1450–1457.

McCormick, C., Duncan, G., Goutsos, K.T., and Tufaro, F. (2000). The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. Proc. Natl. Acad. Sci. *97*, 668–673.

McCourt, P.A.G. (1999). How does the hyaluronan scrap-yard operate? Matrix Biol. *18*, 427–432.

McGovern, R.E., Fernandes, H., Khan, A.R., Power, N.P., and Crowley, P.B. (2012). Protein camouflage in cytochrome c-calixarene complexes. Nat. Chem. *4*, 527–533.

McKenzie, E., Tyson, K., Stamps, A., Smith, P., Turner, P., Barry, R., Hircock, M., Patel, S., Barry, E., Stubberfield, C., et al. (2000). Cloning and expression profiling of Hpa2, a novel mammalian heparanase family member. Biochem. Biophys. Res. Commun. *276*, 1170–1177.

Meen, A.J., Øynebråten, I., Reine, T.M., Duelli, A., Svennevig, K., Pejler, G., Jenssen, T., and Kolset, S.O. (2011). Serglycin is a major proteoglycan in polarized human endothelial cells and is implicated in the secretion of the chemokine GROalpha/CXCL1. J. Biol. Chem. *286*, 2636–2647.

Mehl, E., and Jatzkewitz, H. (1968). Cerebroside 3-sulfate as a physiological substrate of arylsulfatase A. Biochim. Biophys. Acta *151*, 619–627.

Merry, C.L.R., Lyon, M., Deakin, J.A., Hopwood, J.J., and Gallagher, J.T. (1999). Highly Sensitive Sequencing of the Sulfated Domains of Heparan Sulfate. J. Biol. Chem. *274*, 18455–18462.

Meyers, J.R., Planamento, J., Ebrom, P., Krulewitz, N., Wade, E., and Pownall, M.E. (2013). Sulf1 modulates BMP signaling and is required for somite morphogenesis and development of the horizontal myoseptum. Dev. Biol. *378*, 107–121.

Midura, R.J., Calabro, A., Yanagishita, M., and Hascall, V.C. (1995). Nonreducing End Structures of Chondroitin Sulfate Chains on Aggrecan Isolated from Swarm Rat Chondrosarcoma Cultures. J. Biol. Chem. *270*, 8009–8015.

Miech, C., Dierks, T., Selmer, T., Figura, K. von, and Schmidt, B. (1998). Arylsulfatase from Klebsiella pneumoniae Carries a Formylglycine Generated from a Serine. J. Biol. Chem. *273*, 4835–4837.

Mikami, T., and Kitagawa, H. (2013). Biosynthesis and function of chondroitin sulfate. Biochim. Biophys. Acta *1830*, 4719–4733.

Milz, F., Harder, A., Neuhaus, P., Breitkreuz-Korff, O., Walhorn, V., Lübke, T., Anselmetti, D., and Dierks, T. (2013). Cooperation of binding sites at the hydrophilic domain of cell-surface sulfatase Sulf1 allows for dynamic interaction of the enzyme with its substrate heparan sulfate. Biochim. Biophys. Acta BBA - Gen. Subj. *1830*, 5287–5298.

Mondal, S., Roy, D., Camacho-Pereira, J., Khurana, A., Chini, E., Yang, L., Baddour, J., Stilles, K., Padmabandu, S., Leung, S., et al. (2015). HSulf-1 deficiency dictates a metabolic reprograming of glycolysis and TCA cycle in ovarian cancer. Oncotarget *6*, 33705–33719.

Monneau, Y., Arenzana-Seisdedos, F., and Lortat-Jacob, H. (2016). The sweet spot: how GAGs help chemokines guide migrating cells. J. Leukoc. Biol. *99*, 935–953.

Moon, A.F., Edavettal, S.C., Krahn, J.M., Munoz, E.M., Negishi, M., Linhardt, R.J., Liu, J., and Pedersen, L.C. (2004). Structural analysis of the sulfotransferase (3-o-sulfotransferase isoform 3)

involved in the biosynthesis of an entry receptor for herpes simplex virus 1. J. Biol. Chem. *279*, 45185–45193.

Moon, A.F., Xu, Y., Woody, S.M., Krahn, J.M., Linhardt, R.J., Liu, J., and Pedersen, L.C. (2012). Dissecting the substrate recognition of 3-O-sulfotransferase for the biosynthesis of anticoagulant heparin. Proc. Natl. Acad. Sci. U. S. A. *109*, 5265–5270.

Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S., and Rosen, S.D. (2002). Cloning and Characterization of Two Extracellular Heparin-degrading Endosulfatases in Mice and Humans. J. Biol. Chem. *277*, 49175–49185.

Morimoto-Tomita, M., Uchimura, K., Bistrup, A., Lum, D.H., Egeblad, M., Boudreau, N., Werb, Z., and Rosen, S.D. (2005). Sulf-2, a Proangiogenic Heparan Sulfate Endosulfatase, Is Upregulated in Breast Cancer. Neoplasia N. Y. N *7*, 1001–1010.

Morio, H., Honda, Y., Toyoda, H., Nakajima, M., Kurosawa, H., and Shirasawa, T. (2003). EXT gene family member rib-2 is essential for embryonic development and heparan sulfate biosynthesis in Caenorhabditis elegans. Biochem. Biophys. Res. Commun. *301*, 317–323.

Mostovich, L.A., Prudnikova, T.Y., Kondratov, A.G., Gubanova, N.V., Kharchenko, O.A., Kutsenko, O.S., Vavilov, P.V., Haraldson, K., Kashuba, V.I., Ernberg, I., et al. (2012). The TCF4/β-catenin pathway and chromatin structure cooperate to regulate D-glucuronyl C5-epimerase expression in breast cancer. Epigenetics *7*, 930–939.

Moussay, E., Palissot, V., Vallar, L., Poirel, H.A., Wenner, T., El Khoury, V., Aouali, N., Van Moer, K., Leners, B., Bernardin, F., et al. (2010). Determination of genes and microRNAs involved in the resistance to fludarabine in vivo in chronic lymphocytic leukemia. Mol. Cancer *9*, 115.

Muir, H. (1958). The nature of the link between protein and carbohydrate of a chondroitin sulphate complex from hyaline cartilage. Biochem. J. *69*, 195–204.

Myette, J.R., Shriver, Z., Kiziltepe, T., McLean, M.W., Venkataraman, G., and Sasisekharan, R. (2002). Molecular cloning of the heparin/heparan sulfate delta 4,5 unsaturated glycuronidase from Flavobacterium heparinum, its recombinant expression in Escherichia coli, and biochemical determination of its unique substrate specificity. Biochemistry *41*, 7424–7434.

Myette, J.R., Soundararajan, V., Shriver, Z., Raman, R., and Sasisekharan, R. (2009a). Heparin/Heparan Sulfate 6-O-Sulfatase from Flavobacterium heparinum. J. Biol. Chem. *284*, 35177–35188.

Myette, J.R., Soundararajan, V., Behr, J., Shriver, Z., Raman, R., and Sasisekharan, R. (2009b). Heparin/Heparan Sulfate N-Sulfamidase from Flavobacterium heparinum. J. Biol. Chem. *284*, 35189–35200.

Nadanaka, S., Zhou, S., Kagiyama, S., Shoji, N., Sugahara, K., Sugihara, K., Asano, M., and Kitagawa, H. (2013a). EXTL2, a Member of the EXT Family of Tumor Suppressors, Controls Glycosaminoglycan Biosynthesis in a Xylose Kinase-dependent Manner. J. Biol. Chem. *288*, 9321–9333.

Nadanaka, S., Kagiyama, S., and Kitagawa, H. (2013b). Roles of EXTL2, a member of the EXT family of tumour suppressors, in liver injury and regeneration processes. Biochem. J. *454*, 133–145.

Nagai, N., Habuchi, H., Esko, J.D., and Kimata, K. (2004). Stem domains of heparan sulfate 6-O-sulfotransferase are required for Golgi localization, oligomer formation and enzyme activity. J. Cell Sci. *117*, 3331–3341.

Nagai, N., Habuchi, H., Kitazume, S., Toyoda, H., Hashimoto, Y., and Kimata, K. (2007). Regulation of Heparan Sulfate 6-O-Sulfation by β-Secretase Activity. J. Biol. Chem. *282*, 14942–14951.

Nagamine, S., Keino-Masu, K., Shiomi, K., and Masu, M. (2010). Proteolytic cleavage of the rat heparan sulfate 6-O-endosulfatase SulfFP2 by furin-type proprotein convertases. Biochem. Biophys. Res. Commun. *391*, 107–112.

Nagamine, S., Tamba, M., Ishimine, H., Araki, K., Shiomi, K., Okada, T., Ohto, T., Kunita, S., Takahashi, S., Wismans, R.G.P., et al. (2012). Organ-specific Sulfation Patterns of Heparan Sulfate Generated by Extracellular Sulfatases Sulf1 and Sulf2 in Mice. J. Biol. Chem. *287*, 9579–9590.

Nakamura, I., Fernandez-Barrena, M.G., Ortiz-Ruiz, M.C., Almada, L.L., Hu, C., Elsawa, S.F., Mills, L.D., Romecin, P.A., Gulaid, K.H., Moser, C.D., et al. (2013). Activation of the Transcription Factor GLI1 by WNT Signaling Underlies the Role of SULFATASE 2 as a Regulator of Tissue Regeneration. J. Biol. Chem. *288*, 21389–21398.

Narentuya, Takeda-Uchimura, Y., Foyez, T., Zhang, Z., Akama, T.O., Yagi, H., Kato, K., Komatsu, Y., Kadomatsu, K., and Uchimura, K. (2019). GlcNAc6ST3 is a keratan sulfate sulfotransferase for the protein-tyrosine phosphatase PTPRZ in the adult brain. Sci. Rep. *9*.

Narita, K., Staub, J., Chien, J., Meyer, K., Bauer, M., Friedl, A., Ramakrishnan, S., and Shridhar, V. (2006). HSulf-1 Inhibits Angiogenesis and Tumorigenesis In vivo. Cancer Res. *66*, 6025–6032.

Narita, K., Chien, J., Mullany, S.A., Staub, J., Qian, X., Lingle, W.L., and Shridhar, V. (2007). Loss of HSulf-1 Expression Enhances Autocrine Signaling Mediated by Amphiregulin in Breast Cancer. J. Biol. Chem. *282*, 14413–14420.

Nawroth, R., van Zante, A., Cervantes, S., McManus, M., Hebrok, M., and Rosen, S.D. (2007). Extracellular sulfatases, elements of the Wnt signaling pathway, positively regulate growth and tumorigenicity of human pancreatic cancer cells. PloS One *2*, e392.

Neill, T., Schaefer, L., and Iozzo, R.V. (2015). Decoding the Matrix: Instructive Roles of Proteoglycan Receptors. Biochemistry *54*, 4583–4598.

Ohtake-Niimi, S., Kondo, S., Ito, T., Kakehi, S., Ohta, T., Habuchi, H., Kimata, K., and Habuchi, O. (2010). Mice Deficient in N-Acetylgalactosamine 4-Sulfate 6-O-Sulfotransferase Are Unable to Synthesize Chondroitin/Dermatan Sulfate containing N-Acetylgalactosamine 4,6-Bissulfate Residues and Exhibit Decreased Protease Activity in Bone Marrow-derived Mast Cells. J. Biol. Chem. *285*, 20793–20805.

Ono, K., Hattori, H., Takeshita, S., Kurita, A., and Ishihara, M. (1999). Structural features in heparin that interact with VEGF165 and modulate its biological activity. Glycobiology *9*, 705–711.

Otsuki, S., Hanson, S.R., Miyaki, S., Grogan, S.P., Kinoshita, M., Asahara, H., Wong, C.-H., and Lotz, M.K. (2010). Extracellular sulfatases support cartilage homeostasis by regulating BMP and FGF signaling pathways. Proc. Natl. Acad. Sci. U. S. A. *107*, 10202–10207.

Otsuki, S., Alvarez-Garcia, O., Lotz, M.K., and Neo, M. (2019). Role of heparan sulfate 6-0 endosulfatases in intervertebral disc homeostasis. Histol. Histopathol. 18107.

Oustah, A.A., Danesin, C., Khouri-Farah, N., Farreny, M.-A., Escalas, N., Cochard, P., Glise, B., and Soula, C. (2014). Dynamics of Sonic hedgehog signaling in the ventral spinal cord are controlled by intrinsic changes in source cells requiring Sulfatase 1. Development *141*, 1392–1403.

Pacheco, B., Malmström, A., and Maccarana, M. (2009). Two dermatan sulfate epimerases form iduronic acid domains in dermatan sulfate. J. Biol. Chem. *284*, 9788–9795.

Pal, S., Doganges, P.T., and Schubert, M. (1966). The separation of new forms of the proteinpolysaccharides of bovine nasal cartilage. J. Biol. Chem. *241*, 4261–4266.

Pallerla, S.R., Lawrence, R., Lewejohann, L., Pan, Y., Fischer, T., Schlomann, U., Zhang, X., Esko, J.D., and Grobe, K. (2008). Altered Heparan Sulfate Structure in Mice with Deleted NDST3 Gene Function. J. Biol. Chem. *283*, 16885–16894.

Pankonin, M.S., Gallagher, J.T., and Loeb, J.A. (2005). Specific Structural Features of Heparan Sulfate Proteoglycans Potentiate Neuregulin-1 Signaling. J. Biol. Chem. *280*, 383–388.

Pasquato, A., Dettin, M., Basak, A., Gambaretto, R., Tonin, L., Seidah, N.G., and Di Bello, C. (2007). Heparin enhances the furin cleavage of HIV-1 gp160 peptides. FEBS Lett. *581*, 5807–5813.

Patel, V.N., Likar, K.M., Zisman-Rozen, S., Cowherd, S.N., Lassiter, K.S., Sher, I., Yates, E.A., Turnbull, J.E., Ron, D., and Hoffman, M.P. (2008). Specific Heparan Sulfate Structures Modulate FGF10-mediated Submandibular Gland Epithelial Morphogenesis and Differentiation. J. Biol. Chem. *283*, 9308–9317.

Pempe, E.H., Burch, T.C., Law, C.J., and Liu, J. (2012). Substrate specificity of 6-O-endosulfatase (Sulf-2) and its implications in synthesizing anticoagulant heparan sulfate. Glycobiology *22*, 1353–1362.

Peterson, S.M., Iskenderian, A., Cook, L., Romashko, A., Tobin, K., Jones, M., Norton, A., Gómez-Yafal, A., Heartlein, M.W., Concino, M.F., et al. (2010). Human Sulfatase 2 inhibits in vivo tumor growth of MDA-MB-231 human breast cancer xenografts. BMC Cancer *10*, 427.

Petitou, M., Casu, B., and Lindahl, U. (2003). 1976–1983, a critical period in the history of heparin: the discoveryof the antithrombin binding site. Biochimie *85*, 83–89.

Phillips, J.J., Huillard, E., Robinson, A.E., Ward, A., Lum, D.H., Polley, M.-Y., Rosen, S.D., Rowitch, D.H., and Werb, Z. (2012). Heparan sulfate sulfatase SULF2 regulates PDGFRα signaling and growth in human and mouse malignant glioma. J. Clin. Invest. *122*, 911–922.

Pichert, A., Schlorke, D., Franz, S., and Arnhold, J. (2012). Functional aspects of the interaction between interleukin-8 and sulfated glycosaminoglycans. Biomatter *2*, 142–148.

Pikas, D.S., Eriksson, I., and Kjellén, L. (2000). Overexpression of Different Isoforms of Glucosaminyl N-Deacetylase/N-Sulfotransferase Results in Distinct Heparan Sulfate N-Sulfation Patterns. Biochemistry *39*, 4552–4558.

Pinhal, M.A.S., Smith, B., Olson, S., Aikawa, J., Kimata, K., and Esko, J.D. (2001). Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. Proc. Natl. Acad. Sci. *98*, 12984–12989.

Pomin, V.H., and Mulloy, B. (2018). Glycosaminoglycans and Proteoglycans. Pharmaceuticals *11*.

Pratt, T., Conway, C.D., Tian, N.M.M.-L., Price, D.J., and Mason, J.O. (2006). Heparan Sulphation Patterns Generated by Specific Heparan Sulfotransferase Enzymes Direct Distinct Aspects of Retinal Axon Guidance at the Optic Chiasm. J. Neurosci. *26*, 6911–6923.

Préchoux, A., Halimi, C., Simorre, J.-P., Lortat-Jacob, H., and Laguri, C. (2015). C5-Epimerase and 2-O-Sulfotransferase Associate in Vitro to Generate Contiguous Epimerized and 2-O-Sulfated Heparan Sulfate Domains. ACS Chem. Biol. *10*, 1064–1071.

Presto, J., Thuveson, M., Carlsson, P., Busse, M., Wilén, M., Eriksson, I., Kusche-Gullberg, M., and Kjellén, L. (2008). Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. Proc. Natl. Acad. Sci. *105*, 4751–4756.

Proudfoot, A.E.I. (2006). The biological relevance of chemokine–proteoglycan interactions. Biochem. Soc. Trans. *34*, 422–426.

Prudnikova, T.Y., Mostovich, L.A., Kashuba, V.I., Ernberg, I., Zabarovsky, E.R., and Grigorieva, E.V. (2012). miRNA-218 contributes to the regulation of D-glucuronyl C5-epimerase expression in normal and tumor breast tissues. Epigenetics *7*, 1109–1114.

Prydz, K. (2015). Determinants of Glycosaminoglycan (GAG) Structure. Biomolecules *5*, 2003–2022.

Pye, D.A., Vives, R.R., Turnbull, J.E., Hyde, P., and Gallagher, J.T. (1998). Heparan Sulfate Oligosaccharides Require 6-O-Sulfation for Promotion of Basic Fibroblast Growth Factor Mitogenic Activity. J. Biol. Chem. *273*, 22936–22942.

Qin, Y., Ke, J., Gu, X., Fang, J., Wang, W., Cong, Q., Li, J., Tan, J., Brunzelle, J.S., Zhang, C., et al. (2015). Structural and Functional Study of d-Glucuronyl C5-epimerase. J. Biol. Chem. *290*, 4620–4630.

Raman, R., Myette, J.R., Shriver, Z., Pojasek, K., Venkataraman, G., and Sasisekharan, R. (2003). The Heparin/Heparan Sulfate 2-O-Sulfatase from Flavobacterium heparinum A STRUCTURAL AND BIOCHEMICAL STUDY OF THE ENZYME ACTIVE SITE AND SACCHARIDE SUBSTRATE SPECIFICITY. J. Biol. Chem. *278*, 12167–12174.

Ramsbottom, S.A., Maguire, R.J., Fellgett, S.W., and Pownall, M.E. (2014). Sulf1 influences the Shh morphogen gradient during the dorsal ventral patterning of the neural tube in Xenopus tropicalis. Dev. Biol. *391*, 207–218.

Rapraeger, A.C., Krufka, A., and Olwin, B.B. (1991). Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. Science *252*, 1705–1708.

Reijmers, R.M., Vondenhoff, M.F.R., Roozendaal, R., Kuil, A., Li, J.-P., Spaargaren, M., Pals, S.T., and Mebius, R.E. (2010). Impaired Lymphoid Organ Development in Mice Lacking the Heparan Sulfate Modifying Enzyme Glucuronyl C5-Epimerase. J. Immunol. *184*, 3656–3664.

Reijmers, R.M., Groen, R.W.J., Kuil, A., Weijer, K., Kimberley, F.C., Medema, J.P., Kuppevelt, T.H. van, Li, J.-P., Spaargaren, M., and Pals, S.T. (2011). Disruption of heparan sulfate proteoglycan conformation perturbs B-cell maturation and APRIL-mediated plasma cell survival. Blood *117*, 6162–6171.

Reine, T.M., Kusche-Gullberg, M., Feta, A., Jenssen, T., and Kolset, S.O. (2012). Heparan sulfate expression is affected by inflammatory stimuli in primary human endothelial cells. Glycoconj. J. *29*, 67–76.

Rickard, S.M., Mummery, R.S., Mulloy, B., and Rider, C.C. (2003). The binding of human glial cell line-derived neurotrophic factor to heparin and heparan sulfate: importance of 2-O-sulfate groups and effect on its interaction with its receptor, GFRalpha1. Glycobiology *13*, 419–426.

Robinson, C.J., Mulloy, B., Gallagher, J.T., and Stringer, S.E. (2006). VEGF165-binding Sites within Heparan Sulfate Encompass Two Highly Sulfated Domains and Can Be Liberated by K5 Lyase. J. Biol. Chem. *281*, 1731–1740.

Rodén, L., and Smith, R. (1966). Structure of the neutral trisaccharide of the chondroitin 4-sulfate-protein linkage region. J. Biol. Chem. *241*, 5949–5954.

Rong, J., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2000). Expression of heparan sulphate L-iduronyl 2-O-sulphotransferase in human kidney 293 cells results in increased D-glucuronyl 2-O-sulphation. Biochem. J. *346*, 463–468.

Rong, J., Habuchi, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2001). Substrate Specificity of the Heparan Sulfate Hexuronic Acid 2-O-Sulfotransferase. Biochemistry *40*, 5548–5555.

Rosen, S.D., and Lemjabbar-Alaoui, H. (2010). SULF-2: AN EXTRACELLULAR MODULATOR OF CELL SIGNALING AND A CANCER TARGET CANDIDATE. Expert Opin. Ther. Targets *14*, 935–949.

Roy, A.B. (1975). L-ascorbic acid 2-sulphate. A substrate for mammalian arylsulphatases. Biochim. Biophys. Acta *377*, 356–363.

Roy, D., Mondal, S., Wang, C., He, X., Khurana, A., Giri, S., Hoffmann, R., Jung, D.-B., Kim, S.H., Chini, E.N., et al. (2014a). Loss of HSulf-1 promotes altered lipid metabolism in ovarian cancer. Cancer Metab. *2*, 13.

Roy, S., El Hadri, A., Richard, S., Denis, F., Holte, K., Duffner, J., Yu, F., Galcheva-Gargova, Z., Capila, I., Schultes, B., et al. (2014b). Synthesis and biological evaluation of a unique heparin mimetic hexasaccharide for structure-activity relationship studies. J. Med. Chem. *57*, 4511–4520.

Sadir, R., Baleux, F., Grosdidier, A., Imberty, A., and Lortat-Jacob, H. (2001). Characterization of the Stromal Cell-derived Factor-1α-Heparin Complex. J. Biol. Chem. *276*, 8288–8296.

Sadir, R., Imberty, A., Baleux, F., and Lortat-Jacob, H. (2004). Heparan Sulfate/Heparin Oligosaccharides Protect Stromal Cell-derived Factor-1 (SDF-1)/CXCL12 against Proteolysis Induced by CD26/Dipeptidyl Peptidase IV. J. Biol. Chem. *279*, 43854–43860.

Safaiyan, F., Lindahl, U., and Salmivirta, M. (1998). Selective reduction of 6-O-sulfation in heparan sulfate from transformed mammary epithelial cells. Eur. J. Biochem. *252*, 576–582.

Sahota, A.P., and Dhoot, G.K. (2009). A novel SULF1 splice variant inhibits Wnt signalling but enhances angiogenesis by opposing SULF1 activity. Exp. Cell Res. *315*, 2752–2764.

Sarrazin, S., Lamanna, W.C., and Esko, J.D. (2011). Heparan Sulfate Proteoglycans. Cold Spring Harb. Perspect. Biol. *3*.

Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995). A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. Cell *82*, 271–278.

Scholefield, Z., Yates, E.A., Wayne, G., Amour, A., McDowell, W., and Turnbull, J.E. (2003). Heparan sulfate regulates amyloid precursor protein processing by BACE1, the Alzheimer's beta-secretase. J. Cell Biol. *163*, 97–107.

Schumacher, V.A., Schlötzer-Schrehardt, U., Karumanchi, S.A., Shi, X., Zaia, J., Jeruschke, S., Zhang, D., Pavenstädt, H., Pavenstaedt, H., Drenckhan, A., et al. (2011). WT1-dependent sulfatase expression maintains the normal glomerular filtration barrier. J. Am. Soc. Nephrol. JASN *22*, 1286–1296.

Sedita, J., Izvolsky, K., and Cardoso, W.V. (2004). Differential expression of heparan sulfate 6-O-sulfotransferase isoforms in the mouse embryo suggests distinctive roles during organogenesis. Dev. Dyn. *231*, 782–794.

Seffouh, A., Milz, F., Przybylski, C., Laguri, C., Oosterhof, A., Bourcier, S., Sadir, R., Dutkowski, E., Daniel, R., van Kuppevelt, T.H., et al. (2013). HSulf sulfatases catalyze processive and oriented 6-O-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *27*, 2431–2439.

Seffouh, I., Przybylski, C., Seffouh, A., El Masri, R., Vivès, R.R., Gonnet, F., and Daniel, R. (2019). Mass spectrometry analysis of the human endosulfatase Hsulf-2. Biochem. Biophys. Rep. *18*.

Shaw, J.P., Johnson, Z., Borlat, F., Zwahlen, C., Kungl, A., Roulin, K., Harrenga, A., Wells, T.N.C., and Proudfoot, A.E.I. (2004). The X-ray structure of RANTES: heparin-derived disaccharides allows the rational design of chemokine inhibitors. Struct. Lond. Engl. 1993 *12*, 2081–2093.

Sheng, J., Liu, R., Xu, Y., and Liu, J. (2011). The Dominating Role of N-Deacetylase/N-Sulfotransferase 1 in Forming Domain Structures in Heparan Sulfate. J. Biol. Chem. *286*, 19768–19776.

Shipp, E.L., and Hsieh-Wilson, L.C. (2007). Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. Chem. Biol. *14*, 195–208.

Shworak, N.W., Liu, J., Petros, L.M., Zhang, L., Kobayashi, M., Copeland, N.G., Jenkins, N.A., and Rosenberg, R.D. (1999). Multiple Isoforms of Heparan Sulfate d-Glucosaminyl 3-O-Sulfotransferase ISOLATION, CHARACTERIZATION, AND EXPRESSION OF HUMAN cDNAs AND IDENTIFICATION OF DISTINCT GENOMIC LOCI. J. Biol. Chem. *274*, 5170–5184.

Sikora, A.-S., Hellec, C., Carpentier, M., Martinez, P., Delos, M., Denys, A., and Allain, F. (2016). Tumour-necrosis factor-α induces heparan sulfate 6-O-endosulfatase 1 (Sulf-1) expression in fibroblasts. Int. J. Biochem. Cell Biol. *80*, 57–65.

Silbert, J.E., and Sugumaran, G. (1995). Intracellular membranes in the synthesis, transport, and metabolism of proteoglycans. Biochim. Biophys. Acta BBA - Rev. Biomembr. *1241*, 371–384.

Silbert, J.E., and Sugumaran, G. (2002). Biosynthesis of chondroitin/dermatan sulfate. IUBMB Life *54*, 177–186.

Singer, M.S., Phillips, J.J., Lemjabbar-Alaoui, H., Wang, Y.Q., Wu, J., Goldman, R., and Rosen, S.D. (2015). SULF2, a heparan sulfate endosulfatase, is present in the blood of healthy individuals and increases in cirrhosis. Clin. Chim. Acta Int. J. Clin. Chem. *440*, 72–78.

Small, E.M., Sutherland, L., Rajagopalan, K., Wang, S., and Olson, E.N. (2010). MicroRNA-218 regulates vascular patterning by modulation of Slit-Robo signaling. Circ. Res. *107*, 1336–1344.

Smeds, E., Habuchi, H., Do, A.-T., Hjertson, E., Grundberg, H., Kimata, K., Lindahl, U., and Kusche-Gullberg, M. (2003). Substrate specificities of mouse heparan sulphate glucosaminyl 6-O-sulphotransferases. Biochem. J. *372*, 371–380.

Smeds, E., Feta, A., and Kusche-Gullberg, M. (2010). Target selection of heparan sulfate hexuronic acid 2-O-sulfotransferase. Glycobiology *20*, 1274–1282.

Spillmann, D., Witt, D., and Lindahl, U. (1998). Defining the Interleukin-8-binding Domain of Heparan Sulfate. J. Biol. Chem. *273*, 15487–15493.

Stanford, K.I., Bishop, J.R., Foley, E.M., Gonzales, J.C., Niesman, I.R., Witztum, J.L., and Esko, J.D. (2009). Syndecan-1 is the primary heparan sulfate proteoglycan mediating hepatic clearance of triglyceride-rich lipoproteins in mice. J. Clin. Invest. *119*, 3236–3245.

Staples, G.O., Shi, X., and Zaia, J. (2011). Glycomics analysis of mammalian heparan sulfates modified by the human extracellular sulfatase HSulf2. PloS One *6*, e16689.

Staub, J., Chien, J., Pan, Y., Qian, X., Narita, K., Aletti, G., Scheerer, M., Roberts, L.R., Molina, J., and Shridhar, V. (2007). Epigenetic silencing of HSulf-1 in ovarian cancer:implications in chemoresistance. Oncogene *26*, 4969–4978.

Stern, R., and Jedrzejas, M.J. (2006). The Hyaluronidases: Their Genomics, Structures, and Mechanisms of Action. Chem. Rev. *106*, 818–839.

Stickens, D., Zak, B.M., Rougier, N., Esko, J.D., and Werb, Z. (2005). Mice deficient in Ext2 lack heparan sulfate and develop exostoses. Dev. Camb. Engl. *132*, 5055–5068.

Stressler, T., Seitl, I., Kuhn, A., and Fischer, L. (2016). Detection, production, and application of microbial arylsulfatases. Appl. Microbiol. Biotechnol. *100*, 9053–9067.

Stringer, S.E., and Gallagher, J.T. (1997). Specific Binding of the Chemokine Platelet Factor 4 to Heparan Sulfate. J. Biol. Chem. *272*, 20508–20514.

Sugaya, N., Habuchi, H., Nagai, N., Ashikari-Hada, S., and Kimata, K. (2008). 6-O-Sulfation of Heparan Sulfate Differentially Regulates Various Fibroblast Growth Factor-dependent Signalings in Culture. J. Biol. Chem. *283*, 10366–10376.

Sweeney, E.A. (2002). Sulfated polysaccharides increase plasma levels of SDF-1 in monkeys and mice: involvement in mobilization of stem/progenitor cells. Blood *99*, 44–51.

Taghizadeh, E., Kalantar, S.M., Mahdian, R., Sheikhha, M.H., Farashahi-Yazd, E., Ghasemi, S., and Shahbazi, Z. (2015). SULF 1 gene polymorphism, rs6990375 is in significant association with fetus failure in IVF technique. Iran. J. Reprod. Med. *13*, 215–220.

Takashima, Y., Keino-Masu, K., Yashiro, H., Hara, S., Suzuki, T., van Kuppevelt, T.H., Masu, M., and Nagata, M. (2016). Heparan sulfate 6-O-endosulfatases, Sulf1 and Sulf2, regulate glomerular integrity by modulating growth factor signaling. Am. J. Physiol. Renal Physiol. *310*, F395-408.

Tang, R., and Rosen, S.D. (2009). Functional Consequences of the Subdomain Organization of the Sulfs. J. Biol. Chem. *284*, 21505–21514.

Teng, Y.H.-F., Aquino, R.S., and Park, P.W. (2012). Molecular functions of syndecan-1 in disease. Matrix Biol. J. Int. Soc. Matrix Biol. *31*, 3–16.

Tessema, M., Yingling, C.M., Thomas, C.L., Klinge, D.M., Bernauer, A.M., Liu, Y., Dacic, S., Siegfried, J.M., Dahlberg, S.E., Schiller, J.H., et al. (2012). SULF2 Methylation is Prognostic for Lung Cancer Survival and Increases Sensitivity to Topoisomerase-I inhibitors via Induction of ISG15. Oncogene *31*, 4107–4116.

Thacker, B.E., Xu, D., Lawrence, R., and Esko, J.D. (2014). Heparan sulfate 3-O-sulfation: A rare modification in search of a function. Matrix Biol. J. Int. Soc. Matrix Biol. *35*, 60–72.

Theocharis, A.D., Seidel, C., Borset, M., Dobra, K., Baykov, V., Labropoulou, V., Kanakis, I., Dalas, E., Karamanos, N.K., Sundan, A., et al. (2006). Serglycin Constitutively Secreted by Myeloma Plasma Cells Is a Potent Inhibitor of Bone Mineralization in Vitro. J. Biol. Chem. *281*, 35116–35128.

Tran, T.H., Shi, X., Zaia, J., and Ai, X. (2012). Heparan Sulfate 6-O-endosulfatases (Sulfs) Coordinate the Wnt Signaling Pathways to Regulate Myoblast Fusion during Skeletal Muscle Regeneration. J. Biol. Chem. *287*, 32651–32664.

Tsai, T.-T., Ho, N.Y.-J., Fang, H.-C., Lai, P.-L., Niu, C.-C., Chen, L.-H., Chen, W.-J., and Pang, J.-H.S. (2015). Increased sulfatase 1 gene expression in degenerative intervertebral disc cells. J. Orthop. Res. Off. Publ. Orthop. Res. Soc. *33*, 312–317.

Turnbull, J.E., Fernig, D.G., Ke, Y., Wilkinson, M.C., and Gallagher, J.T. (1992). Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate. J. Biol. Chem. *267*, 10337–10341.

Uchimura, K., Kadomatsu, K., Nishimura, H., Muramatsu, H., Nakamura, E., Kurosawa, N., Habuchi, O., El-Fasakhany, F.M., Yoshikai, Y., and Muramatsu, T. (2002). Functional Analysis of the Chondroitin 6-Sulfotransferase Gene in Relation to Lymphocyte Subpopulations, Brain Development, and Oversulfated Chondroitin Sulfates. J. Biol. Chem. *277*, 1443–1450.

Uchimura, K., Morimoto-Tomita, M., Bistrup, A., Li, J., Lyon, M., Gallagher, J., Werb, Z., and Rosen, S.D. (2006). HSulf-2, an extracellular endoglucosamine-6-sulfatase, selectively mobilizes heparin-bound growth factors and chemokines: effects on VEGF, FGF-1, and SDF-1. BMC Biochem. *7*, 2.

Ueno, M., Yamada, S., Zako, M., Bernfield, M., and Sugahara, K. (2001). Structural Characterization of Heparan Sulfate and Chondroitin Sulfate of Syndecan-1 Purified from Normal Murine Mammary Gland Epithelial Cells COMMON PHOSPHORYLATION OF XYLOSE AND DIFFERENTIAL SULFATION OF GALACTOSE IN THE PROTEIN LINKAGE REGION TETRASACCHARIDE SEQUENCE. J. Biol. Chem. *276*, 29134–29140.

Ulmer, J.E., Vilén, E.M., Namburi, R.B., Benjdia, A., Beneteau, J., Malleron, A., Bonnaffé, D., Driguez, P.-A., Descroix, K., Lassalle, G., et al. (2014). Characterization of glycosaminoglycan (GAG) sulfatases from the human gut symbiont Bacteroides thetaiotaomicron reveals the first GAG-specific bacterial endosulfatase. J. Biol. Chem. *289*, 24289–24303.

Vivès, R.R., Pye, D.A., Salmivirta, M., Hopwood, J.J., Lindahl, U., and Gallagher, J.T. (1999). Sequence analysis of heparan sulphate and heparin oligosaccharides. Biochem. J. *339*, 767–773.

Vivès, R.R., Crublet, E., Andrieu, J.-P., Gagnon, J., Rousselle, P., and Lortat-Jacob, H. (2004). A novel strategy for defining critical amino acid residues involved in protein/glycosaminoglycan interactions. J. Biol. Chem. *279*, 54327–54333.

Vivès, R.R., Seffouh, A., and Lortat-Jacob, H. (2014). Post-Synthetic Regulation of HS Structure: The Yin and Yang of the Sulfs in Cancer. Front. Oncol. *3*.

Viviano, B.L., Paine-Saunders, S., Gasiunas, N., Gallagher, J., and Saunders, S. (2004). Domain-specific Modification of Heparan Sulfate by Qsulf1 Modulates the Binding of the Bone Morphogenetic Protein Antagonist Noggin. J. Biol. Chem. *279*, 5604–5611.

Vlodavsky, I., Gross-Cohen, M., Weissmann, M., Ilan, N., and Sanderson, R.D. (2018). Opposing Functions of Heparanase-1 and Heparanase-2 in Cancer Progression. Trends Biochem. Sci. *43*, 18–31.

Wagner, L., Yang, O.O., Garcia-Zepeda, E.A., Ge, Y., Kalams, S.A., Walker, B.D., Pasternack, M.S., and Luster, A.D. (1998). Beta-chemokines are released from HIV-1-specific cytolytic T-cell granules complexed to proteoglycans. Nature *391*, 908–911.

Waldow, A., Schmidt, B., Dierks, T., Bülow, R. von, and Figura, K. von (1999). Amino Acid Residues Forming the Active Site of Arylsulfatase A ROLE IN CATALYTIC ACTIVITY AND SUBSTRATE BINDING. J. Biol. Chem. *274*, 12284–12288.

Walhorn, V., Möller, A.-K., Bartz, C., Dierks, T., and Anselmetti, D. (2018). Exploring the Sulfatase 1 Catch Bond Free Energy Landscape using Jarzynski's Equality. Sci. Rep. *8*, 16849.

Wang, L., Brown, J.R., Varki, A., and Esko, J.D. (2002). Heparin's anti-inflammatory effects require glucosamine 6-O-sulfation and are mediated by blockade of L- and P-selectins. J. Clin. Invest. *110*, 127–136.

Wang, L., Xie, L., Wang, J., Shen, J., and Liu, B. (2013). Correlation between the methylation of SULF2 and WRN promoter and the irinotecan chemosensitivity in gastric cancer. BMC Gastroenterol. *13*, 173.

Wang, S., Ai, X., Freeman, S.D., Pownall, M.E., Lu, Q., Kessler, D.S., and Emerson, C.P. (2004). QSulf1, a heparan sulfate 6-O-endosulfatase, inhibits fibroblast growth factor signaling in mesoderm induction and angiogenesis. Proc. Natl. Acad. Sci. U. S. A. *101*, 4833–4838.

Wei, G., Bai, X., Bame, K.J., Koshy, T.I., Spear, P.G., and Esko, J.D. (2000). Location of the glucuronosyltransferase domain in the heparan sulfate copolymerase (EXT1) by analysis of Chinese hamster ovary cell mutants. J. Biol. Chem.

Weigel, P.H., Hascall, V.C., and Tammi, M. (1997). Hyaluronan Synthases. J. Biol. Chem. *272*, 13997–14000.

Weissmann, M., Arvatz, G., Horowitz, N., Feld, S., Naroditsky, I., Zhang, Y., Ng, M., Hammond, E., Nevo, E., Vlodavsky, I., et al. (2016). Heparanase-neutralizing antibodies attenuate lymphoma tumor growth and metastasis. Proc. Natl. Acad. Sci. U. S. A. *113*, 704–709.

Weissmann, M., Bhattacharya, U., Feld, S., Hammond, E., Ilan, N., and Vlodavsky, I. (2018). The heparanase inhibitor PG545 is a potent anti-lymphoma drug: Mode of action. Matrix Biol. J. Int. Soc. Matrix Biol.

Weyers, A., Yang, B., Solakyildirim, K., Yee, V., Li, L., Zhang, F., and Linhardt, R.J. (2013). Isolation of bovine corneal keratan sulfate and its growth factor and morphogen binding. FEBS J. *280*, 2285–2293.

Whitelock, J.M., Melrose, J., and Iozzo, R.V. (2008). Diverse Cell Signaling Events Modulated by Perlecan. Biochemistry *47*, 11174–11183.

Wijnhoven, T.J.M., Lensen, J.F.M., Rops, A.L.W.M.M., van der Vlag, J., Kolset, S.O., Bangstad, H.-J., Pfeffer, P., van den Hoven, M.J.W., Berden, J.H.M., van den Heuvel, L.P.W.J., et al. (2006). Aberrant heparan sulfate profile in the human diabetic kidney offers new clues for therapeutic glycomimetics. Am. J. Kidney Dis. Off. J. Natl. Kidney Found. *48*, 250–261.

Winterbottom, E.F., and Pownall, M.E. (2009). Complementary expression of HSPG 6-O-endosulfatases and 6-O-sulfotransferase in the hindbrain of Xenopus laevis. Gene Expr. Patterns GEP *9*, 166–172.

Wu, L., Viola, C.M., Brzozowski, A.M., and Davies, G.J. (2015). Structural characterization of human heparanase reveals insights into substrate recognition. Nat. Struct. Mol. Biol. *22*, 1016–1022.

Wuyts, W., and Van Hul, W. (2000). Molecular basis of multiple exostoses: mutations in the EXT1 and EXT2 genes. Hum. Mutat. *15*, 220–227.

Xia, G., Chen, J., Tiwari, V., Ju, W., Li, J.-P., Malmström, A., Shukla, D., and Liu, J. (2002). Heparan Sulfate 3-O-Sulfotransferase Isoform 5 Generates Both an Antithrombin-binding Site and an Entry Receptor for Herpes Simplex Virus, Type 1. J. Biol. Chem. *277*, 37912–37919.

Xu, D., Moon, A.F., Song, D., Pedersen, L.C., and Liu, J. (2008). Engineering sulfotransferases to modify heparan sulfate. Nat. Chem. Biol. *4*, 200–202.

Xu, G., Ji, W., Su, Y., Xu, Y., Yan, Y., Shen, S., Li, X., Sun, B., Qian, H., Chen, L., et al. (2014). Sulfatase 1 (hSulf-1) reverses basic fibroblast growth factor-stimulated signaling and inhibits growth of hepatocellular carcinoma in animal model. Oncotarget *5*, 5029–5039.

Xu, R., Ori, A., Rudd, T.R., Uniewicz, K.A., Ahmed, Y.A., Guimond, S.E., Skidmore, M.A., Siligardi, G., Yates, E.A., and Fernig, D.G. (2012). Diversification of the Structural Determinants of Fibroblast Growth Factor-Heparin Interactions IMPLICATIONS FOR BINDING SPECIFICITY. J. Biol. Chem. *287*, 40061–40073.

Yan, D., and Lin, X. (2009). Shaping Morphogen Gradients by Proteoglycans. Cold Spring Harb. Perspect. Biol. *1*, a002493.

Yang, Y., Yaccoby, S., Liu, W., Langford, J.K., Pumphrey, C.Y., Theus, A., Epstein, J., and Sanderson, R.D. (2002). Soluble syndecan-1 promotes growth of myeloma tumors in vivo. Blood *100*, 610–617.

Yang, Y., Macleod, V., Miao, H.-Q., Theus, A., Zhan, F., Shaughnessy, J.D., Sawyer, J., Li, J.-P., Zcharia, E., Vlodavsky, I., et al. (2007). Heparanase enhances syndecan-1 shedding: a novel mechanism for stimulation of tumor growth and metastasis. J. Biol. Chem. *282*, 13326–13333.

Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P., and Ornitz, D.M. (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. Cell *64*, 841–848.

Yue, X. (2017). Epithelial Deletion of Sulf2 Exacerbates Bleomycin-Induced Lung Injury, Inflammation, and Mortality. Am. J. Respir. Cell Mol. Biol. *57*, 560–569.

Yue, X., Li, X., Nguyen, H.T., Chin, D.R., Sullivan, D.E., and Lasky, J.A. (2008). Transforming Growth Factor-β1 Induces Heparan Sulfate 6-O-Endosulfatase 1 Expression in Vitro and in Vivo. J. Biol. Chem. *283*, 20397–20407.

Yue, X., Shan, B., and Lasky, J.A. (2010). TGF-β: Titan of Lung Fibrogenesis. Curr. Enzyme Inhib. *6*.

Yue, X., Lu, J., Auduong, L., Sides, M.D., and Lasky, J.A. (2013). Overexpression of Sulf2 in idiopathic pulmonary fibrosis. Glycobiology *23*, 709–719.

Zahraei, M., Sheikhha, M.H., Kalantar, S.M., Ghasemi, N., Jahaninejad, T., Rajabi, S., and Mohammadpour, H. (2014). The association of arylendosulfatase 1 (SULF1) gene polymorphism with recurrent miscarriage. J. Assist. Reprod. Genet. *31*, 157–161.

Zak, B.M., Crawford, B.E., and Esko, J.D. (2002). Hereditary multiple exostoses and heparan sulfate polymerization. Biochim. Biophys. Acta BBA - Gen. Subj. *1573*, 346–355.

Zhang, H., Newman, D.R., and Sannes, P.L. (2012a). HSULF-1 inhibits ERK and AKT signaling and decreases cell viability in vitro in human lung epithelial cells. Respir. Res. *13*, 69.

Zhang, L., David, G., and Esko, J.D. (1995). Repetitive Ser-Gly Sequences Enhance Heparan Sulfate Assembly in Proteoglycans. J. Biol. Chem. *270*, 27127–27135.

Zhang, L., Beeler, D.L., Lawrence, R., Lech, M., Liu, J., Davis, J.C., Shriver, Z., Sasisekharan, R., and Rosenberg, R.D. (2001). 6-O-Sulfotransferase-1 Represents a Critical Enzyme in the Anticoagulant Heparan Sulfate Biosynthetic Pathway. J. Biol. Chem. *276*, 42311–42321.

Zhang, L., Yang, M., Yang, D., Cavey, G., Davidson, P., and Gibson, G. (2010). Molecular Interactions of MMP-13 C-Terminal Domain with Chondrocyte Proteins. Connect. Tissue Res. *51*, 230–239.

Zhang, S., Condac, E., Qiu, H., Jiang, J., Gutierrez-Sanchez, G., Bergmann, C., Handel, T., and Wang, L. (2012b). Heparin-induced leukocytosis requires 6-O-sulfation and is caused by blockade of selectin- and CXCL12 protein-mediated leukocyte trafficking in mice. J. Biol. Chem. *287*, 5542–5553.

Zhang, X., Wang, F., and Sheng, J. (2016). “Coding” and “Decoding”: hypothesis for the regulatory mechanism involved in heparan sulfate biosynthesis. Carbohydr. Res. *428*, 1–7.

Zhang, X., Wang, Y., Xie, M., Corbett, C., Singhal, S., Dai, B., Wang, J., Ding, Q., Lu, Q., and Wang, Y. (2018). Downregulating Heparanase-Induced Vascular Normalization: A New Approach To Increase the Bioavailability of Chemotherapeutics in Solid Tumors. Mol. Pharm. *15*, 4303–4309.

Zhao, W., Sala-Newby, G.B., and Dhoot, G.K. (2006). Sulf1 expression pattern and its role in cartilage and joint development. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *235*, 3327–3335.

ZHU, C., HE, L., ZHOU, X., NIE, X., and GU, Y. (2016). Sulfatase 2 promotes breast cancer progression through regulating some tumor-related factors. Oncol. Rep. *35*, 1318–1328.

Zlotnik, A., and Yoshie, O. (2012). The Chemokine Superfamily Revisited. Immunity *36*, 705–716.

Zong, F., Fthenou, E., Wolmer, N., Hollósi, P., Kovalszky, I., Szilák, L., Mogler, C., Nilsonne, G., Tzanakakis, G., and Dobra, K. (2009). Syndecan-1 and FGF-2, but Not FGF Receptor-1, Share a Common Transport Route and Co-Localize with Heparanase in the Nuclei of Mesenchymal Tumor Cells. PLoS ONE *4*.

Résumé

Les Héparanes Sulfates (HS) sont de polysaccharides complexes impliqués dans de nombreux processus biologiques. La structure des HS est contrôlée à la surface cellulaire par une famille particulière d'endosulfatases extracellulaires, les Sulfs. Les Sulfs modifient dramatiquement les propriétés fonctionnelles des HS et sont impliqués dans de nombreux processus physiopathologiques, notamment le cancer. Ces enzymes se composent de deux domaines: un domaine catalytique (CAT) contenant le site actif et un domaine basique hydrophile (HD) responsable de la liaison aux HS. Le but de mon projet de thèse est de caractériser les propriétés structurales et fonctionnelles de la forme humaine HSulf-2, qui demeure à ce jour très mal connues. Dans ce cadre, nous avons tout d'abord étudié les mécanismes de reconnaissance enzyme/substrat et caractérisé deux nouveaux motifs de reconnaissance des HS sur ces enzymes, responsable de leur activité. En utilisant des oligosaccharides naturels et synthétiques, nous avons aussi démontré que le domaine HD n'est pas essentiel pour la reconnaissance des HS, mais permet une désulfatation processive et orientée du polysaccharide. De plus, nous avons identifié un tétrasaccharide comme étant la taille oligosaccharidique minimale requise pour l'activité de HSulf-2. Nos résultats nous ont permis de proposer un nouveau modèle décrivant le processus de désulfatation du HS par HSulf-2. D'autre part, nous avons montré que HSulf-2 est un protéoglycane, car il contient une modification post-traductionnelle unique (chaîne CS de Chondoitin Sulfate) sur son domaine HD. Cette chaîne diminue l'activité enzymatique et la liaison aux HS *in vitro*. Dans le microenvironnement tumoral, en utilisant un modèle de tumeur mammaire orthotopique murin, nous avons montré que la chaîne CS est libérée par protéolyse, conduisant à l'activation de HSulf-2, augmentant la capacité des tumeurs à se développer et à se transformer en métastase. Finalement, nous avons réalisé une étude structurale des Sulfs. Nous avons choisi d'étudier séparément les deux domaines (CAT et HD). Des essais de cristallogenèse ont été menés pour le domaine CAT afin de résoudre sa structure par cristallographie aux rayons X, mais n'ont pu aboutir. En ce qui concerne le HD, nous avons mis en place un protocole de production et de purification de HD d'une manière recombinante et nous avons initiés une étude par RMN ainsi que d'autres techniques biophysiques afin de caractériser structuralement le domaine et d'identifier les sites de liaison aux HS. Nos résultats préliminaires suggèrent que la HD est un domaine non structuré, à l'exception de ses parties N- et C-terminales. L'ensemble de ces travaux devrait nous

permettre de mieux comprendre ces importants mécanismes de régulation des HS et de d'envisager de nouvelles stratégies anticancéreuses ciblant les Sulfs.

Summary

Heparan Sulfate (HS) are complex polysaccharides involved in many biological processes. The structure of HS is regulated at the cell surface by unique extracellular endosulfatases, the Sulfs. Sulfs dramatically change HS functional properties, thereby being implicated in many physiopathological processes including cancer. Sulfs features two domains: a catalytic domain (CAT) that comprises the active site, and an hydrophilic basic domain (HD) responsible for HS binding. The aim of my PhD project is to characterize the structural and the functional properties of the human for HSulf-2, which remains poorly understood. In this context, we have first studied the enzyme/substrate recognition mechanisms. We identified two novel HS binding motifs on these enzymes implicated in their activity. In addition, using natural and synthetic oligosaccharides, we demonstrated that the HD is not essential for HS recognition, but is directs the processive and orientated desulfation of the polysaccharide. Moreover, we showed that a tetrasaccharide is the minimal oligosaccharide size required for HSulf-2 activity. Our results enabled us to propose a new model depicting the desulfation process of HS by the Sulfs. Second, we have shown that HSulf-2 is a proteoglycan, given that it harbors a unique PTM (Chondroitin Sulfate, CS chain) on its HD domain. This chain decreases enzyme activity and HS binding *in vitro*. In the tumoral microenvironment, using a murine orthotropic mammary tumor model, we showed that the CS chain is lost by proteolytic processing, leading to the activation of HSulf-2, and the promotion of tumor growth, vascularization and metastasis. Finally, we have undertaken the structural characterization of the Sulfs. For this, we decided to study separately the two domains found in these enzymes (CAT and HD). Crystallogenesis assays were undertaken for the CAT domain to solve its structure by X-ray crystallography, but were unsuccessful. Regarding the HD, we set up a protocol of production and purification of recombinant HD and we initiated NMR studies and other biophysics analyses in order to structurally characterize the domain and to identify the HS binding sites. Our preliminary results suggest that the HD is an unstructured domain, except for its N- and C-terminal parts. Overall, our data provide significant insights into this critical regulatory step of HS function.